

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2002, 02:52:20 ; Search time 51 Seconds  
(without alignments)  
857.670 Million cell updates/sec

Title: US-09-825-751A-20  
Perfect score: 2447  
Sequence: 1 MDTPRVLLSAVELISFLWDL.....GNVVYKQYEDMVVESCGR 455

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1798	73.5	436	2 B55452	cartilage-derived
2	927.5	37.9	501	2 JC2347	growth/differentia
3	919.5	37.6	495	2 S43294	bone morphogenetic
4	918.5	37.5	501	2 A53452	cartilage-derived
5	691	28.2	125	2 S43295	bone morphogenetic
6	549	22.4	151	2 S43296	bone morphogenetic
7	470	19.2	398	2 JH0687	bone morphogenetic
8	464	19.0	398	2 JH0688	bone morphogenetic
9	453	18.5	405	2 I50608	bone morphogenetic
10	441.5	18.0	408	1 BMRU4	bone morphogenetic
11	436.5	17.8	420	2 I49541	bone morphogenetic
12	434	17.7	393	2 S37073	bone morphogenetic
13	432.5	17.7	408	2 S38343	bone morphogenetic
14	431.5	17.6	396	1 BMRU2	bone morphogenetic
15	429.5	17.6	461	2 S52408	SPDVR1 protein - s
16	426.5	17.4	408	2 JH0801	bone morphogenetic
17	423	17.3	394	2 S45355	bone morphogenetic
18	419.5	17.1	408	2 S58791	bone morphogenetic
19	418.5	17.1	452	2 I49542	bone morphogenetic
20	417.5	17.1	353	2 I50607	bone morphogenetic
21	416.5	17.0	402	2 A45056	osteoegenic protein
22	416.5	17.0	454	1 BMRU5	bone morphogenetic
23	413.5	16.9	401	2 JH0689	bone morphogenetic
24	408.5	16.7	400	2 A49147	bone morphogenetic
25	407	16.6	426	2 JH0690	bone morphogenetic
26	401.5	16.4	431	1 BMRU7	bone morphogenetic
27	401	16.4	430	2 JQ1184	osteoegenic protein
28	395.5	16.2	588	2 A26158	decapentaplegic pr
29	394.5	16.1	510	2 A54798	Vg-1-related prote

30	394.5	16.1	513	1 BMRU6	bone morphogenetic
31	391	16.0	313	2 I51284	bone morphogenetic
32	377	15.4	354	2 S29718	gene nodal protein
33	372	15.2	427	2 A40735	TGF beta homolog d
34	351	14.3	476	2 JC4646	bone morphogenetic
35	350.5	14.3	478	2 JC4838	bone morphogenetic
36	343	14.0	365	2 T43286	cet-1 protein - Ca
37	343	14.0	455	2 A43918	TGF-beta-related p
38	334	13.6	207	2 S37618	vgr protein - rat
39	325.5	13.3	472	1 BMRU3	bone morphogenetic
40	324.5	13.3	360	2 A29619	Vg1 embryonic grow
41	319.5	13.1	372	2 C39564	GDF-1 embryonic gr
42	318	13.0	426	1 B24248	inhibin beta-A cha
43	313.5	12.8	425	2 I47072	inhibin beta-A cha
44	310	12.7	424	1 WFFCBA	inhibin beta-A cha
45	306.5	12.5	425	1 S50898	inhibin beta-A cha

ALIGNMENTS

RESULT 1

B55452  
cartilage-derived morphogenetic protein 2 precursor - bovine (fragment)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 10-Feb-1995 #sequence\_revision 10-Feb-1995 #text\_change 26-May-2000  
C:Accession: B55452  
R:Chang, S.C.; Hoang, B.; Thomas, J.T.; Vukicevic, S.; Luyten, F.P.; Ryba, N.J.P.; Ko  
J. Biol. Chem. 269, 28227-28234, 1994  
A:Title: Cartilage-derived morphogenetic proteins. New members of the transforming gr  
A:Reference number: A55452; MUID:95050604; PMID:7961761  
A:Accession: B55452  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-436 <CHA>  
A:Cross-references: GB:U13661; NID:9632489; PIDN:AAA61416.1; PID:9632490  
C:Superfamily: inhibin

Query Match 73.5%; Score 1798; DB 2; Length 436;  
Best Local Similarity 80.8%; Pred. No. 8.1e-120;  
Matches 354; Conservative 18; Mismatches 48; Indels 18; Gaps 8;

QY	33	SSAELGSTRKGMRSRKEGKQMRAPRDSADGRE---GGEPPQPRQDEPR---AQOPRAQEP 86
DB	2	ASAELGSAKGMRTKEGRMPRAPRENATAREPLDRQEPPEPPQEEPPQRRPPQPEAREPP 61
QY	87	GRGPRVVPHEYMLSIYRTYSIAEKLGINASFQSSKSANTITSFVDRGLDLSHTPLRRQ 146
DB	62	GRGPRVLPHEYMLSIYRTYSIAEKLGINASFQSSKSANTITSFVDRGLDLSHTPLRRQ 121
QY	147	KYLFVDSMLSDKEELVGAELRLFRQAPSAPWGPAGPLHVQLPCLSPLLDARTLDPOG 206
DB	122	KYLFVDSMLSDKEELVGAELRLFRQAPSAPWGPAGPLHVQLPCLSPLLDARTLDPOG 206
QY	207	APPAGWEVDFVWGLRHQPKQLCLLELRAAW-GELDAGEAEARARGPQQPPPPDLRLSLGF 265
DB	180	APRGWEVDFVWGLRHQPKQLCLLELRAAWGCEGPEAAEAEARTPGPQQPPPPDLRLSLGF 239
QY	266	GRVRPQERALLVVTTRQKNLFAMREQLGSA-EAAGPGAGAGGSW-----PPPS 317
DB	240	GRVRPQERALLVVTTRQKNLFAMREQLGSAEAEARARGPQQPPPPPPPPPPPPPP 299
QY	318	GAPDARPLWSPGRRRRRTAFASRHGKSRKSLKPLHVNFKELGWDWIITAPLE 377
DB	300	GTPDAGLWSPG-RRRTAFASRHGKSRKSLKPLHVNFKELGWDWIITAPLE 358
QY	378	YEAYHCEGVCDFFLRHGLEPTNHAIITQLMNSMDPGSTPPSCCVPTKLTPISTIYIDAGN 437
DB	359	YEAYHCEGVCDFFLRHGLEPTNHAIITQLMNSMDPGSTPPSCCVPTKLTPISTIYIDAGN 418
QY	438	NVYKQYEDMVVESCGR 455
DB	419	NVYNEVEEMVESCGR 436

```

Query Match      37.5%; Score 918.5; DB 2; Length 501;
Best Local Similarity 44.3%; Pred. No. 1.5e-57;
Matches 207; Conservative 64; Mismatches 101; Indels 95; Gaps 15;

QY 38 GSTKGMSRREKGMQRPRDSAGREGQEPPRPQDEPRAQQ----- 79
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

C>Date: 20-Oct-1994 #sequence\_revision 10-Nov-1995 #text\_change 19-May-2000  
C/Accession: S43296  
R/Storm, E.E.; Huynh, T.V.; Copeland, N.G.; Jenkins, N.A.; Kingsley, D.M.; Lee, S.J.  
Nature 368, 639-643, 1994  
A>Title: Limb alterations in brachypodism mice due to mutations in a new member of the  
A/Reference number: S43294; MUID:94195427; PMID:8145850  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-151 <STO>  
A/Cross-references: GB:U08339; NID:g488465; PIDN:AAA18780.1; PID:g488466  
C/Superfamily: Inhibin

Query Match 22.4%; Score 549; DB 2; Length 151;  
Best Local Similarity 59.5%; Pred. No. 5.9e-32;  
Matches 103; Conservative 13; Mismatches 35; Indels 22; Gaps 1;

QY 283 RSQRKNLFAEMREOLGSAEAGPCAGACGSGWPPSGAPDARPWLPSPGRRRRRTAFASRH 342  
DB 1 RRRRTALAGTRAGGSGGGGGGGGGGGGGGGG----- 38

QY 343 GKRHGKRSRCKKPLHVNFKELGWDWIIAPLEYAYHCEGVCDPLRSHLEPTNHAI 402  
DB 39 GRGHRGRGRSCRKSHUHFKELGWDWIIAPUDYBAYHCEGVCDPLRSHLEPTNHAI 98

QY 403 IOTLMSNDGCGSTPPSCVCTKTPISILYIDAGNNVYKQYEDMVVYEGCGCR 455  
DB 99 IOTLLNSMAPDAAPASCVPARLSPISILYIDAGNNVYKQYEDMVVYEGCGCR 151

RESULT 7  
JH0687  
bone morphogenetic protein 2I precursor - African clawed frog  
C/Species: Xenopus laevis (African clawed frog)  
C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Jul-1999  
C/Accession: JH0687; S16244  
R/Nishimatsu, S.; Suzuki, A.; Shoda, A.; Murakami, K.; Ueno, N.  
Biochem. Biophys. Res. Commun. 186, 1487-1495, 1992  
A>Title: Genes for bone morphogenetic proteins are differentially transcribed in earl  
A/Reference number: JH0687; MUID:92378616; PMID:1510675  
A/Accession: JH0687  
A/Molecule type: mRNA  
A/Residues: 1-398 <NIS>  
A/Cross-references: GB:X63424; NID:g64585; PIDN:CAA45018.1; PID:g64586  
A/Experimental source: oocyte  
R/Plessow, S.; Koester, M.; Knoechel, W.  
Biochim. Biophys. Acta 1089, 280-282, 1991  
A>Title: cDNA sequence of Xenopus laevis bone morphogenetic protein 2 (BMP-2).  
A/Reference number: S16244; MUID:91274367; PMID:2054389  
A/Accession: S16244  
A/Molecule type: mRNA  
A/Residues: 1-6,'S',8-15,'V',17-232,'N',234-398 <PLE>  
A/Cross-references: EMBL:X55031; NID:g64581; PIDN:CAA38850.1; PID:g64582  
C/Superfamily: Inhibin  
C/Keywords: dimer; glycoprotein  
F/285-398/Product: bone morphogenetic protein 2I #status predicted <MAT>  
F/137,202,340/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.2%; Score 470; DB 2; Length 398;  
Best Local Similarity 31.1%; Pred. No. 6.4e-26;  
Matches 125; Conservative 60; Mismatches 117; Indels 100; Gaps 17;

QY 86 PGRGPRVPVPEHYMLSIYRTY--STAELKLGINASFQSSKS--ANTITSFVDRGLDLSH 140  
DB 65 PTPGKNVIPPYMLDLVHLHLAQLADEGTSAMDQMERASRANTVRSFHE--ESMEE 122

QY 141 TPLRR----OKYLFDSVMSLSDKBEYGAELRFRQAPSAWPGPAGPLH-VOLFPCLSP 195  
DB 123 IPESREKTIORFENLSSIFNEELVTSALRIREFQVEFFESDSSKLHINIYIVKPA 182

QY 196 -----LLDARTLDPOGAPPAGWVEFDVWQGL-----RHQPKQCLCLELRAAWGEL 240  
DB 183 AAASRGVWRLDTRLVHHN---ESKWESEFDVTPATARTWIAKHOPNHGFVYVT----H 235

[illegible]



C:Superfamily: inhibin  
C:Keywords: bone; glycoprotein  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-292/Domain: propeptide #status predicted <PRO>  
F:293-408/Product: bone morphogenetic protein 4 #status predicted <MAT>  
F:143,208,350,365/Binding site: carbohydrate (Asn) #status predicted

Query Match 18.0%; Score 441.5; DB 1; Length 408;  
Best Local Similarity 28.0%; Pred. No. 6.8e-24;  
Matches 135; Conservative 70; Mismatches 164; Indels 113; Gaps 19;

QY 7 LLSAVFLISFLWDLPGGQQASISSCSAEALGSKTK-MRSRKEGMORAPRDSDA---GR 62  
DQ :| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 7 MLMVLLCOVL--LGGASHASLIPTGKKVAETIQGHAGGRSQSHELLRFDEATLIQM 64  
QY 63 EGQEPPQRPODEPRAQQPRAQEPGRGVVPVPHYMILSVIRTYIAEK-----LGINAS 116  
DQ :| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 65 FGLRRFPQPSK-----AVIP-DYMRDLYLQSGEEEEEQIHSTGLEYP 107  
QY 117 FFQSXSANTITSF-VDRGLDDLSTPLRQ-KYLFDFVSMLSDKEELVGAEELFRQAAPS 174  
DQ :| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 108 ERPASR-INTVRSPFHBEHLENIPGTSNSAFRLFNLSIPENEVISAELRLFRE-Q 164  
QY 175 APWGP-PAGPLHVQLPCLSP-----LLLDAITLDPOCAPPAWEVDVWGQL 221  
DQ || | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 165 VDQGPDMERGEFHRINIYEVMKPAAEYVPGHLITRLTLRLVHHN---VTRWETEDV---- 217  
QY 222 RHQPWKQLCELRAAWGELDAGEAARAGPQQPPPDRLSLGFGRVRRVPPQERALLVVVF 281  
DQ || | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 218 -----SPAVALR---WTREKQPNYGIAIEVTH 240  
QY 282 TRSQRKLFAMREQOLGSAAAGPAGAEGSWPPSGADARPWLSPGR-----RR 333  
DQ :| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 241 LHQTERTHOQHVRISRSLPOGS-----GNW-----AQLRPLLVTFGHDGRGHALTTR 287  
QY 334 RRTAFASRGKRHKCKSRKCRSKPLHNFKELGWDDWIATPLEYAYHEGVCDFPLRS 393  
DQ || | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 288 KRAKRSPKHHSQARKKNKNCRRHSXLVDFSDVGWNNDIVAPPGYQAFYCHGDCFFPLAD 347  
QY 394 HLEPTNHAIOTLNMSMDPGTTPSCCVPTKLTPISILYIDAGNNVVKDYEDMVVESCG 453  
DQ || | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 348 HLNSTNHAIOTLVNSYN-SIPIKACCCVPELSAISMLYIDEYDKVVLKNYQEMVVEGCG 406  
QY 454 CR 455  
DQ ||  
DB 407 CR 408

RESULT 11  
I49541  
bone morphogenetic protein 4 - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 02-Jul-1996 #sequence.revision 02-Jul-1996 #text\_change 16-Jul-1999  
C:Accession: I49541; S29523; B34201  
R:Feng, J.Q.; Chen, D.; Cooney, A.J.; Tsai, M.; Harris, M.A.; Tsai, S.Y.; Feng, M.; Mund  
J. Biol. Chem. 270, 28364-28373, 1995  
A>Title: The mouse bone morphogenetic protein-4(BMP4) gene: Analysis of promoter utiliza  
A:Reference number: I49541; MUID:96081880; PMID:7499338  
A:Accession: I49541  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-420 <RES>  
A:Cross-references: GB:I47480; NID:g994733; PIDN:AAC37698.1; PID:g994734  
R:Dickinson, M.E.; van der Meer-de Jong, R.; Hogan, B.L.M.  
submitted to the EMBL Data Library, December 1990  
A:Description: Nucleotide sequence of the mouse Bone Morphogenetic Protein-4 (BMP-4) cDN  
A:Reference number: S29523  
A:Accession: S29523  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 13-420 <DIC>  
A:Cross-references: EMBL:X56848; NID:g50180; PIDN:CAA40179.1; PID:g50181  
R:Dickinson, M.E.; Kobrin, M.S.; Silan, C.M.; Kingsley, D.M.; Justice, M.J.; Miller, D.A.

peland, N.G.; Jenkins, N.A.  
Genomics 6, 505-520, 1990  
A>Title: Chromosomal localization of seven members of the murine TGF-beta superfamily  
A:Reference number: A34201; MUID:90228966; PMID:1970330  
A:Accession: B34201  
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual  
A:Molecule type: mRNA  
A:Residues: 253-420 <DI2>  
C:Genetics:  
A:Gene: BMP-4  
A:Introns: 11/1; 137/1  
C:Superfamily: inhibin

Query Match 17.8%; Score 436.5; DB 2; Length 420;  
Best Local Similarity 28.5%; Pred. No. 1.6e-23;  
Matches 139; Conservative 62; Mismatches 163; Indels 123; Gaps 19;

QY 7 LLSAVLIFISFWDLPFGQAASISSCSAEGLGPKG-MRSRKEGKMQRAPRSDA----GR 62  
:  
DB 19 MLMVLLVCQVL--LGGASHASLIPECTKKKVAETIQGHAGRRSQSHELLRFDEATILQM 76  
:  
QY 63 EGQEPQRPODEPRAOQPRQAEPGRGVPRVPVPHYMLSIYRTYSIAEK-----LGINASF 117  
:  
DB 77 FGLRRRPQPSKS-----AVIP-DYMRDLRYLKSGSEEEEEESQSOGTGLEY 119  
:  
QY 118 FQSSKS-ANTITTSFVDRLDDLSHTPLRRQ-----KYLFDVSMLSDKELVGAELRLF-- 170  
:  
DB 120 PERPASRANIVSRPHRE--EHELENIPTGSSESAFRFLFNLSIPENEVISSAELRLFREQ 177  
:  
QY 171 --QAPSAPWG-----PPAG--PLHVQLFPCLSPILLDDARTLDPOCAPPAWEVF 215  
:  
DB 178 VDQGFDWEQGFHRINIYEVMKPPAEMVPGH-----LITRLDTRLVHHN---VTRWETF 228  
:  
QY 216 DVWQGLRHQPWKQLCLELRAAWGCDELDAEAERARQQPPPDRLSLGFGRRVRPQPER 275  
:  
DB 229 DV-----SPAVLR-----WTREKQPNYGL 247  
:  
QY 276 ALLVVFTRSQRKNLFAEMREQLGSAEAGPCAGAEGSWPPPSGADARPWLPSGR---- 331  
:  
DB 248 AIEVTLHQTHOGQHVRISRSLPQSGDW-----QLRPLLVTFFGHDGRG 294  
:  
QY 332 ---RRRTAFASRGHKRGHKSKRLCRSKPLHNFKELGWDDWTIIAPLEYAYHCEGVCD 388  
:  
DB 295 HTLTRRRAKRSPKHHPQRSKRNNKNCRRHSILYDFSDVGWNWDIVAPPGYQAFYCHGDGP 354  
:  
QY 389 FPLRSHLEPTNHAIOTLMNSMDPGSTPPSCCVPTKLTPISILIYIDAGNNVVYKQIEDMV 448  
:  
DB 355 FPLADHLNSTNHAIQTVLVNSVN--SSIPACCCVPTLSAISMLYLDEYDKVKVLNYQEMV 413  
:  
QY 449 VESCGCR 455  
:  
DB 414 VEGCGCR 420

RESULT 12  
S37073  
bone morphogenetic protein 2 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 16-Jul-1999  
C:Accession: S37073  
R:Feng, J.Q.; Chen, D.; Feng, M.; Harris, M.A.; Mundy, G.R.; Harris, S.E.  
submitted to the EMBL Data Library, September 1993  
A:Description: cDNA sequence of fetal rat calvarial osteoblast bone morphogenetic pro  
A:Reference number: S37073  
A:Accession: S37073  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-393 <FEN>  
A:Cross-references: EMBL:Z25868; NID:g397950; PIDN:CAA81088.1; PID:g397951  
C:Superfamily: inhibin

Query Match 17.7%; Score 434; DB 2; Length 393;  
Best Local Similarity 30.2%; Pred. No. 2.2e-23;



```
QY 202 LDPQAGPAGWEVDFWQGLRHQPWKQLCLELRAANGELDAGAEARARGPQPPPPDLR 261
Db 197 VN-QNA--SRWESFDYTPAMR--W-----TAQCHANHGFTVVEVAHLEKQ----- 237
QY 262 SLGFGRRVRPPQGERALLVVFTRSOKNLFAEMREQLGSAEAGPGAGAGSMPPPSGAPD 321
Db 238 --GVSKR-----HVRISLSLHD-----EHSW-----SQ 259
QY 322 ARPWLSPG-----RRRRTAFASRHGRHKKSRLL--CSKKPLHVNFKELGWD 370
Db 260 IRPLLVTFGHDGKGPLHKKRQA-----KHKQRKRLKSSCKRHPPLYVDFSDVGW 312
QY 371 WIAPLEYEAYHCEGVCDFPLSHLEPTNHAIQTLMNSMDPGSTPPSCCVTKLPISI 430
Db 313 WIVAPPGYHAFYCHGECPPFLADHLNSTNHAIVQTLVNSVN-SKIPKACCVPTLSAISM 371
QY 431 LYIDAGNNVVKQYEDMVVESCGR 455
Db 372 LYIDENEKVVLKYNQDMVVEGCGCR 396

RESULT 15
S52408
SPDVR1 protein - sea urchin (Strongylocentrotus purpuratus)
C:Species: Strongylocentrotus purpuratus (purple urchin)
C:Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 20-Sep-1999
C:Accession: S52408
R:Ponce, M.R.; Micol, J.L.; Davidson, E.H.
submitted to the EMBL Data Library, February 1995
A:Description: SPDVR1, a member of the transforming growth factor-beta superfamily expre
A:Reference number: S52408
A:Accession: S52408
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-461 <PON>
A:Cross-references: EMBL:248313; NID:g673496; PID:g673497
C:Superfamily: inhibin

Query Match 17.6%; Score 429.5; DB 2; Length 461;
Best Local Similarity 27.7%; Pred. No. 5.5e-23;
Matches 140; Conservative 68; Mismatches 177; Indels 121; Gaps 19;

QY 6 VLLSAVFLISFLWDL---PGFQOASISSCCSAELGSKGMRKKGKMQRAPRDSAGR 62
Db 19 LILSLFFGPGLSWDFYSGDEQDLSLARERRAANYNPSPHMTWERNETIQ-----EILNI 74
QY 63 EGQEPQPRP-----ODEPRAQ-----QPRAQEPGCGPRVPVPHYMLSIYRTYSTA 108
Db 75 LGLQHRPRPPLSRGGNQFCQFTWSYRTLNIDEQSGHPSETEPQPGGLASNAIYN 134
QY 109 EKLGINA-----SFFQSSKSANTITS---FVDRGLDLSHTPLRRQKVLFDVS 153
Db 135 DSSGIGSVMSGTVFNVTREYQAVSQADTMSLPVHYKDAAIETDEH-----RYRFDIG 188
QY 154 MLSDEELVGAELRFRQAPSAPWPPAGPLHVQLF-----PCLSPLLDARTLDPOGA 207
Db 189 RIPOGETVTSAEALRVFRDA--GROGRSLYRIDVLLRERSGDSRSPVYLDSTIV---CA 243
QY 208 PPAGWEVDV-----WOLRHPQWKQLCLELRAANGELDAGEAEARARGPQPPPPDLR 261
Db 244 GDHGLVFDMTSATSTW---RSYPGANVGLQURVE--SLQGLNID-----PTDAG 288
QY 262 SLGFGRRVRPPQGERALLVVFTRSQ-----RKNLFAEMREQLGSAEAGPGAGAGS 312
Db 289 VYGVGN--NEGPEPWVFFQNEEVATNSHLRNRRAATROKKG----- 332
QY 313 WPPPSGAPDARWLPSPGRRRR---TAFASRHGRHKKSRLLRCSKKPLHVNFKELGW 368
Db 333 -----GKRPRKPTDNDIASR--DSASSLNSDWCKRKNLFVNFEEDLDW 374
QY 369 DWTIAPLEYEAYHCEGVCDFPLSHLEPTNHAIQTLMNSMDPGSTPPSCCVTKLP 428
Db 375 QEWIIAPLGYAFYCGGCAFFPLNGHANATNHAIVQTLVHMHSPSHVPQPCCAPTKLSPI 434
```

```
QY 429 SILYIDAGNNVVKQYEDMVVESCGR 454
Db 435 TVLYYDDSRNVVLKYYKKNMVVRACGC 460
```

Search completed: November 25, 2002, 02:58:59  
Job time : 52 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2002, 02:54:00 ; Search time 36 seconds  
(without alignments)  
371.873 Million cell updates/sec

Title: US-09-825-751A-20

Perfect score: 2447

Sequence: 1 MDTPRVLLSAVFLISFLWDL.....GNVVYKQVDMVWSCGR 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

- 1: /cgn2.6/protdata/2/iaa/5A\_COMB.pep:\*\*
- 2: /cgn2.6/protdata/2/iaa/5B\_COMB.pep:\*\*
- 3: /cgn2.6/protdata/2/iaa/6A\_COMB.pep:\*\*
- 4: /cgn2.6/protdata/2/iaa/6B\_COMB.pep:\*\*
- 5: /cgn2.6/protdata/2/iaa/PCTUS\_COMB.pep:\*\*
- 6: /cgn2.6/protdata/2/iaa/backfiles1.pep:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1747	71.4	321	1	US-08-362-670B-26
2	1747	71.4	321	3	US-08-333-576C-26
3	1747	71.4	321	4	US-08-808-324-26
4	1747	71.4	321	5	PCT-US94-14030A-26
5	1202.5	49.1	263	1	US-08-362-670B-32
6	1202.5	49.1	263	3	US-08-333-576C-32
7	1202.5	49.1	263	4	US-08-808-324-32
8	1202.5	49.1	263	5	PCT-US94-14030A-32
9	927.5	37.9	501	2	US-08-288-508C-2
10	919.5	37.6	495	1	US-08-455-559-10
11	919.5	37.6	495	4	US-09-145-060-10
12	919.5	37.6	495	5	PCT-US94-00657-10
13	917.5	37.5	401	3	US-08-289-222E-3
14	917.5	37.5	401	4	US-09-054-528B-3
15	765	31.3	388	1	US-08-362-670B-34
16	765	31.3	388	3	US-08-333-576C-34
17	765	31.3	388	4	US-08-808-324-34
18	765	31.3	388	5	PCT-US94-14030A-34
19	731	29.9	134	1	US-08-581-529B-6
20	731	29.9	134	3	US-09-097-616-6
21	731	29.9	134	5	PCT-US94-07762-6
22	691.5	28.3	294	1	US-08-362-670B-2
23	691.5	28.3	294	3	US-08-333-576C-2
24	691.5	28.3	294	4	US-08-808-324-2
25	691.5	28.3	294	5	PCT-US94-14030A-2
26	661	27.0	119	1	US-08-581-529B-7
27	661	27.0	119	3	US-09-097-616-7

28	661	27.0	119	5	PCT-US94-07762-7	Sequence 7, Appli
29	627	25.6	411	1	US-08-362-670B-28	Sequence 28, Appl
30	627	25.6	411	3	US-08-333-576C-28	Sequence 28, Appl
31	627	25.6	411	4	US-08-808-324-28	Sequence 28, Appl
32	627	25.6	411	5	PCT-US94-14030A-28	Sequence 28, Appl
33	618	25.3	240	1	US-08-362-670B-30	Sequence 30, Appl
34	618	25.3	240	3	US-08-333-576C-30	Sequence 30, Appl
35	618	25.3	240	4	US-08-808-324-30	Sequence 30, Appl
36	618	25.3	240	5	PCT-US94-14030A-30	Sequence 30, Appl
37	553	22.6	161	2	US-08-581-528A-6	Sequence 6, Appli
38	553	22.6	161	5	PCT-US94-07799-6	Sequence 6, Appli
39	549	22.4	119	1	US-08-455-559-13	Sequence 13, Appl
40	549	22.4	119	4	US-09-145-060-13	Sequence 13, Appl
41	549	22.4	119	5	PCT-US94-00657-13	Sequence 13, Appl
42	549	22.4	120	1	US-08-362-670B-4	Sequence 4, Appli
43	549	22.4	120	3	US-08-333-576C-4	Sequence 4, Appli
44	549	22.4	120	4	US-08-808-324-4	Sequence 4, Appli
45	549	22.4	120	5	PCT-US94-14030A-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1  
US-08-362-670B-26  
; Sequence 26, Application US/08362670B  
; Patent No. 5658882  
; GENERAL INFORMATION:  
; APPLICANT: Celeste, Anthony J.  
; APPLICANT: Wozney, John  
; APPLICANT: Rosen, Vicki A.  
; APPLICANT: Wolfman, Neil  
; APPLICANT: Thomsen, Gerald H.  
; APPLICANT: Melton, Douglas A.  
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENETICS INSTITUTE, INC.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/362,670B  
; FILING DATE: December 22, 1994  
; CLASSIFICATION: 51A  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lazar, Steven R.  
; REGISTRATION NUMBER: 32,618  
; REFERENCE/DOCKET NUMBER: 5202-D  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617 498-8260  
; TELEFAX: 617 876-5851  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 321 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-362-670B-26

Query Match 71.4%; Score 1747; DB 1; Length 321;  
Best Local Similarity 100.0%; Pred. No. 4.3e-141;  
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 DLSHTPLRROKYLFDYSMLSDKEELVGAELRFRQAPSPAGPLHVQLFPCLSPLL 196  
|||||

Db 3 DLSHTPLRRQKYLFDVMSLSDKEELVGAELRLFRQAPSAPWGPAGPLHVQLFPCLSPLL 62  
QY 197 LDARTLDPOGAPAGWEVFDVWQGLRHQPKQLCLELRAWGLDAGEAEARAGPQQPP 256  
Db 63 LDARTLDPOGAPAGWEVFDVWQGLRHQPKQLCLELRAWGLDAGEAEARAGPQQPP 122  
QY 257 PDLRLSLGFRGVRVPPQERALLVVFTRRSQKNLFAEMRQOLGSAEAGPGAGAGSWPPP 316  
Db 123 PDLRLSLGFRGVRVPPQERALLVVFTRRSQKNLFAEMRQOLGSAEAGPGAGAGSWPPP 182  
QY 317 SGAPDARWLPSPGRRRRRTAFASRHGKRGKSLRCLSKPLHVNFKELGWDWIIAPL 376  
Db 183 SGAPDARWLPSPGRRRRRTAFASRHGKRGKSLRCLSKPLHVNFKELGWDWIIAPL 242  
QY 377 EYEAVHCEGVCDFPLRSHLEPTNHAIQIOTLMSMDPGSTPPSCCVPTKLTPTISILYIDAG 436  
Db 243 EYEAVHCEGVCDFPLRSHLEPTNHAIQIOTLMSMDPGSTPPSCCVPTKLTPTISILYIDAG 302  
QY 437 NNVYKQYEDMVESGCR 455  
Db 303 NNVYKQYEDMVESGCR 321

RESULT 2  
US-08-333-576C-26  
; Sequence 26, Application US/08333576C  
; Patent No. 6027919  
; GENERAL INFORMATION:  
; APPLICANT: Celeste, Anthony J.  
; APPLICANT: Wozney, John  
; APPLICANT: Rosen, Vicki A.  
; APPLICANT: Wolfman, Neil  
; APPLICANT: Thomsen, Gerald H.  
; APPLICANT: Melton, Douglas A.  
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENETICS INSTITUTE, INC.  
; STREET: 87 Cambridgepark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/333.576C  
; FILING DATE: No. 6027919ember 2, 1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lazar, Steven R.  
; REGISTRATION NUMBER: 32,618  
; REFERENCE/DOCKET NUMBER: 5202-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617 498-8260  
; TELEFAX: 617 876-5851  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 321 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-333-576C-26

Query Match 71.4%; Score 1747; DB 3; Length 321;  
Best Local Similarity 100.0%; Pred. No. 4.3e-141;  
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 DLSHTPLRRQKYLFDVMSLSDKEELVGAELRLFRQAPSAPWGPAGPLHVQLFPCLSPLL 196  
|||||

Db 3 DLSHTPLRRQKYLFDVMSLSDKEELVGAELRLFRQAPSAPWGPAGPLHVQLFPCLSPLL 62  
QY 197 LDARTLDPOGAPAGWEVFDVWQGLRHQPKQLCLELRAWGLDAGEAEARAGPQQPP 256  
Db 63 LDARTLDPOGAPAGWEVFDVWQGLRHQPKQLCLELRAWGLDAGEAEARAGPQQPP 122  
QY 257 PDLRLSLGFRGVRVPPQERALLVVFTRRSQKNLFAEMRQOLGSAEAGPGAGAGSWPPP 316  
Db 123 PDLRLSLGFRGVRVPPQERALLVVFTRRSQKNLFAEMRQOLGSAEAGPGAGAGSWPPP 182  
QY 317 SGAPDARWLPSPGRRRRRTAFASRHGKRGKSLRCLSKPLHVNFKELGWDWIIAPL 376  
Db 183 SGAPDARWLPSPGRRRRRTAFASRHGKRGKSLRCLSKPLHVNFKELGWDWIIAPL 242  
QY 377 EYEAVHCEGVCDFPLRSHLEPTNHAIQIOTLMSMDPGSTPPSCCVPTKLTPTISILYIDAG 436  
Db 243 EYEAVHCEGVCDFPLRSHLEPTNHAIQIOTLMSMDPGSTPPSCCVPTKLTPTISILYIDAG 302  
QY 437 NNVYKQYEDMVESGCR 455  
Db 303 NNVYKQYEDMVESGCR 321

RESULT 3  
US-08-808-324-26  
; Sequence 26, Application US/08808324  
; Patent No. 6284872  
; GENERAL INFORMATION:  
; APPLICANT: Celeste, Anthony J.  
; APPLICANT: Wozney, John  
; APPLICANT: Rosen, Vicki A.  
; APPLICANT: Wolfman, Neil  
; APPLICANT: Thomsen, Gerald H.  
; APPLICANT: Melton, Douglas A.  
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENETICS INSTITUTE, INC.  
; STREET: 87 Cambridgepark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/808.324  
; FILING DATE: Herewith  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lazar, Steven R.  
; REGISTRATION NUMBER: 32,618  
; REFERENCE/DOCKET NUMBER: 5202-D  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617 498-8260  
; TELEFAX: 617 876-5851  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 321 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-808-324-26

Query Match 71.4%; Score 1747; DB 4; Length 321;  
Best Local Similarity 100.0%; Pred. No. 4.3e-141;  
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 DLSHTPLRRQKYLFDVMSLSDKEELVGAELRLFRQAPSAPWGPAGPLHVQLFPCLSPLL 196  
|||||

Db 3 DLSTPLRRQKYLFDYSMLSDKEELVGAELRLFRQAPSPWGPAGPLHVQLPCLSPLL 62  
QY 197 LDARTLDPQAGPAGVEFDVWQGLRHQPKQLCLELRAAWGELDAGEAARARGPOPPP 256  
Db 63 LDARTLDPQAGPAGVEFDVWQGLRHQPKQLCLELRAAWGELDAGEAARARGPOPPP 122  
QY 257 PDLRLSGFGRRRVPQERALLVVFTRSQKKNLFAEMREQLGSAEAGPAGAGSGSWPPP 316  
Db 123 PDLRLSGFGRRRVPQERALLVVFTRSQKKNLFAEMREQLGSAEAGPAGAGSGSWPPP 182  
QY 317 SGAPDARPLSPGRRRRRTAFASRGKRGKSKRLCSKKPLHVNFKELGWDWIIAPL 376  
Db 183 SGAPDARPLSPGRRRRRTAFASRGKRGKSKRLCSKKPLHVNFKELGWDWIIAPL 242  
QY 377 EYAYHCEGVCDPLRSHLEPTNHAIQTLMNSMDPGSTPPSCCVPTKLTPIISILYIDAG 436  
Db 243 EYAYHCEGVCDPLRSHLEPTNHAIQTLMNSMDPGSTPPSCCVPTKLTPIISILYIDAG 302  
QY 437 NNVYKQYEDMVVESCGR 455  
Db 303 NNVYKQYEDMVVESCGR 321

## RESULT 4

PCT-US94-14030A-26  
; Sequence 26, Application PC/TUS9414030A  
; GENERAL INFORMATION:  
; APPLICANT: GENETICS INSTITUTE, INC.  
; APPLICANT: PRESIDENT AND FELLOWS OF HARVARD COLLEGE  
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENETICS INSTITUTE, INC.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: PCT/US94/14030A  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/164,103  
; FILING DATE: 07-DEC-1993  
; APPLICATION NUMBER: US 08/217,780  
; FILING DATE: 25-MAR-1994  
; APPLICATION NUMBER: US 08/333,576  
; FILING DATE: 02-NOV-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lazar, Steven R.  
; REGISTRATION NUMBER: 32,618  
; REFERENCE/DOCKET NUMBER: 5202D-PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617 498-8260  
; TELEFAX: 617 876-5851  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 321 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US94-14030A-26

Query Match 71.4%; Score 1747; DB 5; Length 321;  
Best Local Similarity 100.0%; Pred. No. 4.3e-141;  
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 DLSTPLRRQKYLFDYSMLSDKEELVGAELRLFRQAPSPWGPAGPLHVQLPCLSPLL 196  
Db 3 DLSTPLRRQKYLFDYSMLSDKEELVGAELRLFRQAPSPWGPAGPLHVQLPCLSPLL 62  
QY 197 LDARTLDPQAGPAGVEFDVWQGLRHQPKQLCLELRAAWGELDAGEAARARGPOPPP 256  
Db 63 LDARTLDPQAGPAGVEFDVWQGLRHQPKQLCLELRAAWGELDAGEAARARGPOPPP 122  
QY 257 PDLRLSGFGRRRVPQERALLVVFTRSQKKNLFAEMREQLGSAEAGPAGAGSGSWPPP 316  
Db 123 PDLRLSGFGRRRVPQERALLVVFTRSQKKNLFAEMREQLGSAEAGPAGAGSGSWPPP 182  
QY 317 SGAPDARPLSPGRRRRRTAFASRGKRGKSKRLCSKKPLHVNFKELGWDWIIAPL 376  
Db 183 SGAPDARPLSPGRRRRRTAFASRGKRGKSKRLCSKKPLHVNFKELGWDWIIAPL 242  
QY 377 EYAYHCEGVCDPLRSHLEPTNHAIQTLMNSMDPGSTPPSCCVPTKLTPIISILYIDAG 436  
Db 243 EYAYHCEGVCDPLRSHLEPTNHAIQTLMNSMDPGSTPPSCCVPTKLTPIISILYIDAG 302  
QY 437 NNVYKQYEDMVVESCGR 455  
Db 303 NNVYKQYEDMVVESCGR 321

## RESULT 5

US-08-362-670B-32  
; Sequence 32, Application US/08362670B  
; Patent No. 5658882  
; GENERAL INFORMATION:  
; APPLICANT: Celeste, Anthony J.  
; APPLICANT: Wozney, John  
; APPLICANT: Rosen, Vicki A.  
; APPLICANT: Wolfman, Neil  
; APPLICANT: Thomsen, Gerald H.  
; APPLICANT: Melton, Douglas A.  
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENETICS INSTITUTE, INC.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/362,670B  
; FILING DATE: December 22, 1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lazar, Steven R.  
; REGISTRATION NUMBER: 32,618  
; REFERENCE/DOCKET NUMBER: 5202-D  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617 498-8260  
; TELEFAX: 617 876-5851  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 263 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-362-670B-32

Query Match 49.1%; Score 1202.5; DB 1; Length 263;  
Best Local Similarity 83.3%; Pred. No. 7.7e-95;  
Matches 230; Conservative 6; Mismatches 21; Indels 19; Gaps 4;





QY 300 AEAAGCAGAGSWPPSPGAPDAPWLPSPGRRRRRTAFASRHGKRHGKSRKPL 359  
Db 112 AEA-----AGAEGSCAPSPGSDPTGSLWLPSPGRRRRRTAFASRHGKRHGKSRKPL 167  
QY 360 HVNFKELGWDWIIAPLEYEAYHCEGVCDFPLRSHLEPTNHAIIQTILNMSMDPGSTPPSC 419  
Db 168 HVNFKELGWDWIIAPLEYEAYHCEGVCDFPLRSHLEPTNHAIIQTILNMSMDPGSTPPSC 227  
QY 420 CVPTKLTPIISILYIDAGNNVVYKQYEDMVVESCGR 455  
Db 228 CVPTKLTPIISILYIDAGNNVVYKQYEDMVVESCGR 263

## RESULT 8

PCT-US94-14030A-32  
; Sequence 32, Application PC/TUS9414030A  
; GENERAL INFORMATION:  
; APPLICANT: GENETICS INSTITUTE, INC.  
; APPLICANT: PRESIDENT AND FELLOWS OF HARVARD COLLEGE  
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENETICS INSTITUTE, INC.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/14030A  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/164,103  
; FILING DATE: 07-DEC-1993  
; APPLICATION NUMBER: US 08/217,780  
; FILING DATE: 25-MAR-1994  
; APPLICATION NUMBER: US 08/333,576  
; FILING DATE: 02-NOV-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lazar, Steven R.  
; REGISTRATION NUMBER: 32,618  
; REFERENCE/DOCKET NUMBER: 5202D-PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617 498-8260  
; TELEFAX: 617 876-5851  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 263 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US94-14030A-32

Query Match 49.1%; Score 1202.5; DB 5; Length 263;  
Best Local Similarity 83.3%; Pred. No. 7.7e-95;  
Matches 230; Conservative 6; Mismatches 21; Indels 19; Gaps 4;  
QY 180 PAGPLHVQLPCLSPDLLDARTLDPOGAPGAGWEVDVWGLRHQPWKQLCLELRAWGE 239  
Db 7 PAG-----PVLRSSGTQPR---PAG--KSFVDWVWGLRQPWKQLCLELRAWGE 51  
QY 240 LDAGEAEARAGQQPPPPDLRLSLGFGRRVPPQERALLVVFTRSKNLFAMRQQLGS 299  
Db 52 LDAGDTGARAGQQPPPPDLRLSLGFGRRVPPQERALLVVFTRSKNLFAMRQQLGS 111  
QY 300 AEAAGCAGAGSWPPSPGAPDAPWLPSPGRRRRRTAFASRHGKRHGKSRKPL 359  
Db 112 AEA-----AGAEGSCAPSPGSDPTGSLWLPSPGRRRRRTAFASRHGKRHGKSRKPL 167

Db 112 AEA-----AGAEGSCAPSPGSDPTGSLWLPSPGRRRRRTAFASRHGKRHGKSRKPL 167  
QY 360 HVNFKELGWDWIIAPLEYEAYHCEGVCDFPLRSHLEPTNHAIIQTILNMSMDPGSTPPSC 419  
Db 168 HVNFKELGWDWIIAPLEYEAYHCEGVCDFPLRSHLEPTNHAIIQTILNMSMDPGSTPPSC 227  
QY 420 CVPTKLTPIISILYIDAGNNVVYKQYEDMVVESCGR 455  
Db 228 CVPTKLTPIISILYIDAGNNVVYKQYEDMVVESCGR 263

## RESULT 9

US-08-288-508C-2  
; Sequence 2, Application US/08288508C  
; Patent No. 5994094  
; GENERAL INFORMATION:  
; APPLICANT: H tten, Gertrud  
; APPLICANT: Neidhardt, Helge  
; APPLICANT: Paulista, Michael  
; TITLE OF INVENTION: NEW GROWTH/DIFFERENTIATING FACTOR OF  
; TITLE OF INVENTION: THE TGF- FAMILY  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP  
; STREET: 655 Fifteenth Street N.W. Suite 330  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-5701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/288,508C  
; FILING DATE: 10-AUG-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE P 43 26 829.3  
; FILING DATE: 10-AUG-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE P 44 18 222.8  
; FILING DATE: 25-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE P 44 20 157.5  
; FILING DATE: 09-JUN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: JAHNS, Kristina M.  
; REGISTRATION NUMBER: P-41,092  
; REFERENCE/DOCKET NUMBER: P564-4019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)638-5000  
; TELEFAX: (202)638-4810  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 501 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-288-508C-2

Query Match 37.9%; Score 927.5; DB 2; Length 501;  
Best Local Similarity 44.7%; Pred. No. 4.6e-71;  
Matches 209; Conservative 64; Mismatches 98; Indels 97; Gaps 16;  
QY 38 GSTKGRMSRKEGKQMRAPRDSADAGREGQEPQPRPQDEPRAQQ-----79  
Db 81 GQTGGLTQPKDEPKLIP-----PRPG-GPEPKGHPPTQRTQATARTVTPKGOLPGKAP 134  
QY 80 -----PRAQEPFGGPRVVPHEYMLSIYRTYSIAEKLGINASFFQ 119  
Db 135 PKAGSVPSFLLKKAREPGPPPEPKPEFRPPPTTPHEYMLSYRLTSLDADKGGNSVKL 194





CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-5701  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICANT: POHL, JENS  
FILING DATE: 25-AUG-1999  
CLASSIFICATION: 424  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/289,222  
FILING DATE: 12-AUG-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: DE P 44 23 190.3  
FILING DATE: 07-JUL-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: EPO 92102324.8  
FILING DATE: 12-FEB-1992  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP93/00350  
FILING DATE: 12-FEB-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: KITTS, MONICA CHIN  
REGISTRATION NUMBER: 36,105  
REFERENCE/DOCKET NUMBER: P564-9021  
TELEPHONE: 202/638-5000  
TELEFAX: 202/638-4810  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 401 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-289-222E-3

Query Match 37.5%; Score 917.5; DB 3; Length 401;  
Best Local Similarity 46.0%; Pred. No. 2.4e-70;  
Matches 202; Conservative 59; Mismatches 87; Indels 91; Gaps 14;

QY 67 PQRPPQDEPRAQ-----PRAQPPGR 88  
Db 4 PEPKPGHPQTRQARTVTPKQLPGGKAPKAGSVPSFLLKKAREPGPPREPEPR 63  
QY 89 GRVVPHEYMLSYRTYSAEKLGINASFQSSKANTITSFVDRGLDLSHTPLRQKY 148  
Db 64 PPTPHEYMLSYRTLSADRGKGGNSVKLEAGLANTITSFIDKQDGRGPV-VRKQY 122  
QY 149 LFDVMSLSKEELVGAEHLFRQAPS-----APWGGPAGPLHVLPCLSLD-----PLLLD 198  
Db 123 VFDISAL-ENKDLGALAEHLRLTKKPSDTAKPAAPGGGAAQL--KLSSCPGSRQPASLLD 179  
QY 199 ARTLDPQAGPAGWEVDVMOGLRH-OPWKOLCLELRAANGELDAGEAARAGPQQPP 257  
Db 180 VRVS--PGLDGSGWEVDLWKLFRNFKNSAQLCLEL-AW-----ERGRA-----221  
QY 258 PDLRSLGFRVRPPQERALLVFTTSQRKNL-FAEMREQLGSAEAGPGAGAGSWPPP 316  
Db 222 VDLRGLGFDRAARQVHEKALFLVFGRTKKRDLFFNEIKARSGQDDKT-----268  
QY 317 SGAPDARPLPSPGRRRRRTAFASRKGKRGKSRUCSKKPLHVNFKELGWDWIIAPL 376  
Db 269 -----VYEFLSQ-RRKRRLAPLATROGKRPSKLNKARCSRKALHVNFKDMGWDWIIAPL 322  
QY 377 EYAFHCEGVCDDPLRSHLEPTNHAIIQTLNMSMDPGSTPPSCCVPTKLTPTSLYIDAG 436  
Db 323 EYAFHCEGLCEPFLRSHLEPTNHAIIQTLNMSMDPESTPPTCCVPTSLPISILFDSA 382

QY 437 NNVVYKQYEDMVVESCGR 455  
Db 383 NNVVYKQYEDMVVESCGR 401  
RESULT 14  
US-09-054-526B-3  
Sequence 3, Application US/09054526B  
Patent No. 6197550  
GENERAL INFORMATION:  
APPLICANT: H TTEN, GERTRUD  
APPLICANT: NEIDHARDT, HELGE  
APPLICANT: BECHTOLD, ROLF  
APPLICANT: POHL, JENS  
TITLE OF INVENTION: DNA SEQUENCES ENCODING NOVEL  
TITLE OF INVENTION: GROWTH/DIFFERENTIATION FACTORS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIKAIIDO, MARCELSTEIN, MURRAY & ORAM LLP  
STREET: 655 FIFTEENTH STREET, N. W., G STREET LOBBY,  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-5701  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/054,526B  
FILING DATE: 03-APR-1998  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/289,222  
FILING DATE: 12-AUG-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: DE P 44 23 190.3  
FILING DATE: 01-JUL-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: EPO 92102324.8  
FILING DATE: 12-FEB-1992  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP93/00350  
FILING DATE: 12-FEB-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: KITTS, MONICA CHIN  
REGISTRATION NUMBER: 36,105  
REFERENCE/DOCKET NUMBER: P564-8005  
TELEPHONE: 202/638-5000  
TELEFAX: 202/638-4810  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 401 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-054-526B-3

Query Match 37.5%; Score 917.5; DB 4; Length 401;  
Best Local Similarity 46.0%; Pred. No. 2.4e-70;  
Matches 202; Conservative 59; Mismatches 87; Indels 91; Gaps 14;

QY 67 PQRPPQDEPRAQ-----PRAQPPGR 88  
Db 4 PEPKPGHPQTRQARTVTPKQLPGGKAPKAGSVPSFLLKKAREPGPPREPEPR 63  
QY 89 GRVVPHEYMLSYRTYSAEKLGINASFQSSKANTITSFVDRGLDLSHTPLRQKY 148  
Db 64 PPTPHEYMLSYRTLSADRGKGGNSVKLEAGLANTITSFIDKQDGRGPV-VRKQY 122

QY 149 LFDVSMLSKKEELVGAELRLFRQAPS-----APWGPAGPLHVQLPCLLS-----PULLD 198  
Db 123 VFDISAL-EKQGLGAEELRLKRPDSTAKPAAPGGRAQL--KLSCSPSGRQPASLLD 179  
QY 199 ARTLDPQAGPAGWEVDFWQGLRH-QPWKQLCLELRAAGELDAGEAEARARGPQQPPP 257  
Db 180 VRSV--PGLDGSWEVFDLWKLFERNKSAQLCLELE-AW---ERGRA-----221  
QY 258 PLRSIGFGRVRPPQGERALLVFTVRSQRNL-FAEMREQLGSAEAGPAGAGSWPPP 316  
Db 222 VDLRGLGFDRARQVHEKALFLVFGRTKKRDLEFNEIKARSGDDKT-----268  
QY 317 SGAPDARPLPSPGRRRRRTAFASRCKGHKSKRLRCSKKPLHVNFKELGWDWIIAPL 376  
Db 269 -----VYELFSQ-RKRRAPLATROCKRPSNLKARCSCRKALHVNFKMGWDWIIAPL 322  
QY 377 EYEAHCEGVCDFPLRSHLEPTNHAIQIOLMNSMDPGSTCCVPTKLTPTISILYIDAG 436  
Db 323 EYEAHCEGLCEPLRSHLEPTNHAVIQIOLMNSMDPESTPTCCVTRLSPIISILFDSA 382  
QY 437 NNVYKQYEDMVVESCGR 455  
Db 383 NNVYKQYEDMVVESCGR 401

RESULT 15  
US-08-362-670B-34  
; Sequence 34, Application US/08362670B  
; Patent No. 5658882  
; GENERAL INFORMATION:  
; APPLICANT: Celeste, Anthony J.  
; APPLICANT: Wozney, John  
; APPLICANT: Rosen, Vicki A.  
; APPLICANT: Wolfman, Neil  
; APPLICANT: Thomsen, Gerald H.  
; APPLICANT: Melton, Douglas A.  
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: GENETICS INSTITUTE, INC.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/362,670B  
; FILING DATE: December 22, 1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lazar, Steven R.  
; REGISTRATION NUMBER: 32,618  
; REFERENCE/DOCKET NUMBER: 5202-D  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617 498-8260  
; TELEFAX: 617 876-5851  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 388 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-362-670B-34

Query Match 31.3%; Score 765; DB 1; Length 388;  
Best Local Similarity 51.2%; Pred. No. 2.3e-57;  
Matches 172; Conservative 36; Mismatches 78; Indels 50; Gaps 10;

QY 146 QKYLFDVSMLSKKEELVGAELRLFRQ-APSAPWGPAGPLHVQLPFC-----LSPLLLDAR 200  
Db 77 QSELFDFVSSINDADEVVGAEELRVLRGSPESGPGSWTSPPLLLSTCPGAARAPRLIYSR 136  
QY 201 TLDPOGAPAGWEVDFWQGLRH-----QPWKQLCLELRAAGELDAGEAEARARGPQQP 255  
Db 137 AAEP--LVGQRWEAFDVADAMRRHRREPFRPPAFCLLLRAVAGPV-----P 180  
QY 256 PPDLRLSLGF--GRRVRPPQERALLVFTVRSQRK-NLFAEMREQ---LGSAAEAGPAG 308  
Db 181 SPLALRLGFGWPGGGGSAAEERAVLVVSRQTKESLREIRAQARALGAALASEP---237  
QY 309 AEGSWPPPAGDARWLPSPGRRRRRTAFASRH-----GKRHGKSKRLRCSKKPL 359  
Db 238 -----LPDPCTGTASPRAVITGGRRRRRTALAGTRTAQGGGGAGRGHGRGRSRCSRKPL 292  
QY 360 HVNFKELGWDWIIAPLEVEAHCEGVCDFPLRSHLEPTNHAIQIOLMNSMDPGSTPPSC 419  
Db 293 HVDFKELGWDWIIAPLDYEAHCEGLCDFPLRSHLEPTNHAIQIOLMNSMAPDAAPASC 352  
QY 420 CVPTKLTPTISILYIDAGNNVYKQYEDMVVESCGR 455  
Db 353 CVPARLSPIISILYIDAGNNVYKQYEDMVVESCGR 388

Search completed: November 25, 2002, 02:59:49  
Job time : 37 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2002, 02:55:26 ; Search time 30 Seconds  
(without alignments)  
237.533 Million cell updates/sec

Title: US-09-825-751A-20

Perfect score: 2447

Sequence: 1 MDTPRVLLSAVFLISFLMDL.....GNVVYKQYEDMVVESCGR 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100480 seqs, 15661496 residues

Total number of hits satisfying chosen parameters: 100480

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1798	73.5	436	10	US-09-730-772-14
2	1798	73.5	436	10	US-09-735-849-14
3	1747	71.4	321	9	US-09-945-182-26
4	1202.5	49.1	263	9	US-09-945-182-32
5	927.5	37.9	501	8	US-08-981-490B-1
6	919.5	37.6	495	9	US-09-880-708-10
7	918.5	37.5	502	9	US-09-813-398-37
8	903.5	36.9	501	10	US-09-730-772-13
9	903.5	36.9	501	10	US-09-735-849-13
10	765	31.3	388	9	US-09-945-182-34
11	691.5	28.3	294	9	US-09-945-182-2
12	627	25.6	411	9	US-09-945-182-28
13	618	25.3	240	9	US-09-945-182-30
14	549	22.4	119	8	US-08-945-459A-1
15	549	22.4	119	9	US-08-880-708-13
16	549	22.4	119	9	US-09-068-253-2
17	549	22.4	120	9	US-09-945-182-4
18	442.5	18.1	408	10	US-09-749-728B-69
19	442.5	18.1	409	9	US-09-813-398-27

20	431.5	17.6	396	9	US-10-044-716-2
21	431.5	17.6	396	10	US-09-952-360-2
22	431.5	17.6	397	9	US-09-813-398-24
23	427	17.4	437	10	US-09-784-911-4
24	426	17.4	419	10	US-09-784-911-2
25	422	17.2	427	10	US-09-784-911-6
26	418	17.1	425	9	US-09-813-398-32
27	416.5	17.0	403	9	US-09-813-398-31
28	416.5	17.0	455	9	US-09-813-398-28
29	409.5	16.7	433	10	US-09-784-911-12
30	407.5	16.7	433	10	US-09-784-911-8
31	402.5	16.4	435	10	US-09-784-911-14
32	402.5	16.4	451	10	US-09-784-911-10
33	401.5	16.4	431	8	US-08-822-186-2
34	401.5	16.4	431	8	US-08-937-755-2
35	401.5	16.4	431	9	US-09-982-543A-10
36	401.5	16.4	431	10	US-09-045-331-2
37	401.5	16.4	431	10	US-09-828-607-2
38	400	16.3	72	9	US-09-945-182-13
39	399	16.3	432	10	US-09-361-741-3
40	372	15.2	427	12	US-10-002-278-2
41	359.5	14.7	143	12	US-10-002-278-3
42	357	14.6	72	9	US-09-945-182-15
43	355.5	14.5	356	12	US-10-084-037-1
44	354	14.5	118	9	US-09-859-211-37
45	354	14.5	118	9	US-09-880-708-15

ALIGNMENTS

RESULT 1  
US-09-730-772-14  
; Sequence 14, Application US/09730772  
; Patent No. US2001001131A1  
; GENERAL INFORMATION:  
; APPLICANT: Luyten, Frank P.  
; APPLICANT: Moos, Jr., Malcolm  
; APPLICANT: Chang, Steven Chao-Huan  
; TITLE OF INVENTION: CARTILAGE-DERIVED MORPHOGENETIC  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Knobbe, Martens, Olson & Bear  
; STREET: 620 Newport Center Drive, 16th Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: U.S.A.  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: FASTSEQ for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09730,772  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/836,081  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bartfeld, Neil S  
; REGISTRATION NUMBER: 39,901  
; REFERENCE/DOCKET NUMBER: NIH099.001APC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-235-8550  
; TELEFAX: 619-235-0176  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 436 amino acids  
; TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-730-772-14

Query Match 73.5%; Score 1798; DB 10; Length 436;  
Best Local Similarity 80.8%; Pred. No. 2.6e-120;  
Matches 354; Conservative 18; Mismatches 48; Indels 18; Gaps 8;

QY 33 SSAELGSKGMSRKGMORAPRSDAGRE--GOEPOPQDEPR---AQOPRAQEP 86  
DB 2 ASAEUGSAGKMTREGRMPAPRENATAREPLDQEPPEPQPEPQPEAREPP 61  
QY 87 GRGPRVVPHEYMLSIYRTYSIAEKLGINASFQSSKSANTITSFVDRGLDLSHTPLRRQ 146  
DB 62 GRGPRVVPHEYMLSIYRTYSIAEKLGINASFQSSKSANTITSFVDRGLDLSHTPLRRQ 121  
QY 147 KYLFVMSLKEELVGAELRFRQAPSPAGPAGLVHQLFPCLSPLLLDARTLDPOG 206  
DB 122 KYLFVSTLSKEELVGADVRLFRQAPALAPPAAALRLP-VAPAAAGSAEP-GPAG 179  
QY 207 APPAGWEFVWQGRHOPWKLCLLELRAAW-GELDAGEAEARARGPQPPDLRLSLGF 265  
DB 180 APRGWEFVWGRGURPQWKLCLLELRAAWGEGPAAEDEARTPGQPPPPDLRLSLGF 239  
QY 266 GRVRPPOERALLVFTSRQKNLFAEMREQLGSA-EAAGPGAGAGSW-----PPPS 317  
DB 240 GRVRTPOERALLVFTSRQKNLFAEMREQLGSAEATEVVGPGGAGSGPPPPPPPPPS 299  
QY 318 GAPDAPWLPSPGRRRRRTAFASRGKRGKKSRLRCSKKPLHVNFKELGWDWDWIAPLE 377  
DB 300 GTPDAGLWSPSG-RRRTAFASRGKRGKKSRLRCSKKPLHVNFKELGWDWDWIAPLE 358  
QY 378 YEAYHCEGVCDPPLRSHLEPTNHAIIQTLNMSMDPGSTPPSCCVTKLTPTISILYIDAGN 437  
DB 359 YEAYHCEGVCDPPLRSHLEPTNHAIIQTLNMSMDPGSTPPSCCVTKLTPTISILYIDAGN 418  
QY 438 NVVYKQYEDMVYVSCGCR 455  
DB 419 NVVYNEYEEMVYVSCGCR 436

RESULT 2  
US-09-735-849-14  
Sequence 14, Application US/09735849  
Patent No. US20010037017A1  
GENERAL INFORMATION:  
APPLICANT: Luyten, Frank P.  
APPLICANT: Moos, Jr., Malcolm  
APPLICANT: Chang, Steven Chao-Huan  
TITLE OF INVENTION: CARTILAGE-DERIVED MORPHOGENETIC  
TITLE OF INVENTION: PROTEINS  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 620 Newport Center Drive, 16th Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09735.849  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/836.081  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:

NAME: Bartfeld, Neil S  
REGISTRATION NUMBER: 39,901  
REFERENCE/DOCKET NUMBER: NIH099.001APC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176  
TELEX:  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 436 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-735-849-14

Query Match 73.5%; Score 1798; DB 10; Length 436;  
Best Local Similarity 80.8%; Pred. No. 2.6e-120;  
Matches 354; Conservative 18; Mismatches 48; Indels 18; Gaps 8;

QY 33 SSAELGSKGMSRKGMORAPRSDAGRE--GOEPOPQDEPR---AQOPRAQEP 86  
DB 2 ASAEUGSAGKMTREGRMPAPRENATAREPLDQEPPEPQPEPQPEAREPP 61  
QY 87 GRGPRVVPHEYMLSIYRTYSIAEKLGINASFQSSKSANTITSFVDRGLDLSHTPLRRQ 146  
DB 62 GRGPRVVPHEYMLSIYRTYSIAEKLGINASFQSSKSANTITSFVDRGLDLSHTPLRRQ 121  
QY 147 KYLFVMSLKEELVGAELRFRQAPSPAGPAGLVHQLFPCLSPLLLDARTLDPOG 206  
DB 122 KYLFVSTLSKEELVGADVRLFRQAPALAPPAAALRLP-VAPAAAGSAEP-GPAG 179  
QY 207 APPAGWEFVWQGRHOPWKLCLLELRAAW-GELDAGEAEARARGPQPPDLRLSLGF 265  
DB 180 APRGWEFVWGRGURPQWKLCLLELRAAWGEGPAAEDEARTPGQPPPPDLRLSLGF 239  
QY 266 GRVRPPOERALLVFTSRQKNLFAEMREQLGSA-EAAGPGAGAGSW-----PPPS 317  
DB 240 GRVRTPOERALLVFTSRQKNLFAEMREQLGSAEATEVVGPGGAGSGPPPPPPPPPS 299  
QY 318 GAPDAPWLPSPGRRRRRTAFASRGKRGKKSRLRCSKKPLHVNFKELGWDWDWIAPLE 377  
DB 300 GTPDAGLWSPSG-RRRTAFASRGKRGKKSRLRCSKKPLHVNFKELGWDWDWIAPLE 358  
QY 378 YEAYHCEGVCDPPLRSHLEPTNHAIIQTLNMSMDPGSTPPSCCVTKLTPTISILYIDAGN 437  
DB 359 YEAYHCEGVCDPPLRSHLEPTNHAIIQTLNMSMDPGSTPPSCCVTKLTPTISILYIDAGN 418  
QY 438 NVVYKQYEDMVYVSCGCR 455  
DB 419 NVVYNEYEEMVYVSCGCR 436

RESULT 3  
US-09-945-182-26  
Sequence 26, Application US/09945182  
Patent No. US20020160494A1  
GENERAL INFORMATION:  
APPLICANT: Celeste, Anthony J.  
Wozney, John  
Rosen, Vicki A.  
Wolfman, Neil  
Thomsen, Gerald H.  
Melton, Douglas A.  
TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENETICS INSTITUTE, INC.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02140



COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/09/945,182  
FILING DATE: 31-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/808,324  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Lazar, Steven R.  
REGISTRATION NUMBER: 32,618  
REFERENCE/DOCKET NUMBER: 5202-D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617 498-8260  
TELEFAX: 617 876-5851  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 321 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 26:  
US-09-945-182-26

Query Match 71.4%; Score 1747; DB 9; Length 321;  
Best Local Similarity 100.0%; Pred. No. 7.6e-117;  
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 DLSHTPLRQKYLFDVSMLSKDEELVGAELRLFRQAPSPAGPPLHVLQPCLSPLL 196  
Db 3 DLSHTPLRQKYLFDVSMLSKDEELVGAELRLFRQAPSPAGPPLHVLQPCLSPLL 62

QY 197 LDARTLDPQAGPAGVEFDVWGLRHQPKQLCCLRAAWGELDAEAEARAGFQQPP 256  
Db 63 LDARTLDPQAGPAGVEFDVWGLRHQPKQLCCLRAAWGELDAEAEARAGFQQPP 122

QY 257 PDLRLSGFGRVRRPQERALLVFTSRQKLNFAEMRQLGSAEAAAGPAGAGSGWPPP 316  
Db 123 PDLRLSGFGRVRRPQERALLVFTSRQKLNFAEMRQLGSAEAAAGPAGAGSGWPPP 182

QY 317 SGAPDARPLSPGRRRRRTAFASRHGKRSRCKKPLHVNFKELGWDWIIAPL 376  
Db 183 SGAPDARPLSPGRRRRRTAFASRHGKRSRCKKPLHVNFKELGWDWIIAPL 242

QY 377 EYAYHCEGVCDPPLRSHLEPTNHAIIQTLNMSMDPGSTPPSCCVPTKLTPI SILYIDAG 436  
Db 243 EYAYHCEGVCDPPLRSHLEPTNHAIIQTLNMSMDPGSTPPSCCVPTKLTPI SILYIDAG 302

QY 437 NNVVYKQYEDMVVESCGR 455  
Db 303 NNVVYKQYEDMVVESCGR 321

RESULT 4  
US-09-945-182-32  
Sequence 32, Application US/09945182  
Patent No. US20020160494A1  
GENERAL INFORMATION:  
APPLICANT: Celeste, Anthony J.  
Wozney, John  
Rosen, Vicki A.  
Wolfman, Neil  
Thomsen, Gerald H.  
Melton, Douglas A.  
TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENETICS INSTITUTE, INC.  
STREET: 87 CambridgePark Drive

CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/09/945,182  
FILING DATE: 31-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/808,324  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Lazar, Steven R.  
REGISTRATION NUMBER: 32,618  
REFERENCE/DOCKET NUMBER: 5202-D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617 498-8260  
TELEFAX: 617 876-5851  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 263 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 32:  
US-09-945-182-32

Query Match 49.1%; Score 1202.5; DB 9; Length 263;  
Best Local Similarity 83.3%; Pred. No. 2.1e-78;  
Matches 230; Conservative 6; Mismatches 21; Indels 19; Gaps 4;

QY 180 PAGLHVQLFPCLSPLLDARTLDPQAGPAGVEFDVWGLRHQPKQLCCLRAAWGE 239  
Db 7 PAGLHVQLFPCLSPLLDARTLDPQAGPAGVEFDVWGLRHQPKQLCCLRAAWGE 51

QY 240 LDAGEAEARAGFQQPPDLRLSGFGRVRRPQERALLVFTSRQKLNFAEMRQLG 299  
Db 52 LDAGDTCARAGFQQPPDLRLSGFGRVRRPQERALLVFTSRQKLNFAEMRQLG 111

QY 300 AEAGPAGAGGAGSWPPSGAPDARPLSPGRRRRRTAFASRHGKRSRCKKPL 359  
Db 112 AEA---AGAGGAPAGSGSDTGSWLPSPGRRRRRTAFASRHGKRSRCKKPL 167

QY 360 HVNFKELGWDWIIAPLEYAYHCEGVCDPPLRSHLEPTNHAIIQTLNMSMDPGSTPPSC 419  
Db 168 HVNFKELGWDWIIAPLEYAYHCEGVCDPPLRSHLEPTNHAIIQTLNMSMDPGSTPPSC 227

QY 420 CVPTKLTPI SILYIDAGNNVVYKQYEDMVVESCGR 455  
Db 228 CVPTKLTPI SILYIDAGNNVVYKQYEDMVVESCGR 263

RESULT 5  
US-08-981-490B-1  
Sequence 1, Application US/08981490B  
Patent No. US20020045568A1  
GENERAL INFORMATION:  
APPLICANT: Hotten, Gertrud  
APPLICANT: Pohl, Jens  
APPLICANT: Bechtold, Rolf  
APPLICANT: Paulista, Michael  
APPLICANT: Unsicker, Klaus  
TITLE OF INVENTION: USE OF MP52 OR MP121 FOR TREATING AND PREVENTING DISEASES OF T  
FILE REFERENCE: 100564-07032  
CURRENT APPLICATION NUMBER: US/08/981,490B  
CURRENT FILING DATE: 1998-05-18  
PRIOR APPLICATION NUMBER: PCT/EP96/03065

;  
; PRIOR FILING DATE: 1996-07-12  
; PRIOR APPLICATION NUMBER: DE/195 25 416.3  
; PRIOR FILING DATE: 1995-07-12  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO: 1  
; LENGTH: 501  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-08-981-490B-1

Query Match 37.9%; Score 927.5; DB 8; Length 501;  
Best Local Similarity 44.7%; Pred. No. 1.3e-58;  
Matches 209; Conservative 64; Mismatches 98; Indels 97; Gaps 16;

QY 38 GSPKMRSRKEGKQAPRDSADREGQGPQPPQDEPRAQ-- 79  
Db 81 GQTGGTQPKKDEPKLP-----PRPG-GPEPKPGHPQTROATARTVTPKQLPGGKAP 134  
QY 80 -----PRAQEPGRVVPVPEHYMLSIYRTYSIAEKLGINASFFQ 119  
Db 135 PRAGSVSSFLKKAREPGPPREPKEFPPTPTPHEYMLSLYRTLSADRGKGNSSVKL 194  
QY 120 SSKSANTITSFVDRGLDDLSHTPLRQKYLFDVSMDSKEELVGAELRLFRQAPS----- 174  
Db 195 EAGLANITISFIDKGQDDRGPV--VRKQRYVFDISAL--EKDGLLGAELRLIRLKKPSDTAKP 252  
QY 175 -APWGPAGPLHVQLPCLLS---PLLLDARTLDPGAPAGWEVDVQGLRH-QPWKQ 228  
Db 253 APGGRGAQL--KLSSCSGQAPSLLDVRSV--PGLDGSQGEVFDIWKLFNFKNLSAQ 308  
QY 229 LCLELRAANGELDAGAEARAGPQPPPPDLKSLGFGRRVRPPOBRALLVVFTRSORKN 288  
Db 309 LCLELE-AW---ERGRA-----VDLRLGLGFDRAARQVHEKALFLVFGRTKKRD 352  
QY 289 L-FAEMREQLGSAAAGPAGAGEGSWPPSPGADARPWLPSGRRRRRTAFASRHGKRHG 347  
Db 353 LFFNEIKARSGQDDKT-----VVEYLFQ--RRKRRAPLATRQGRKPS 393  
QY 348 KKSRLCSKKPLHVNFKELGWDWIITAPLEYEAYHCEGVCDPFLRSHLEPTNHAIIQTLM 407  
Db 394 KNLKARCSKKALHVNFKDGMWDWIITAPLEYEAFHCEGLCEFFLRSHLEPTNHAIIQTLM 453  
QY 408 NSMDPSTPPSCVPTKLPISILYIDAGNNVYKYEDMVVESCGR 455  
Db 454 NSMDPESTPTCCVPTRLSPISILFIDSANNVYKYEDMVVESCGR 501

RESULT 6  
US-09-880-708-10  
; Sequence 10, Application US/09880708  
; Patent No. US20020165361A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, Se-Jin  
; Huynh, Thanh  
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-5  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Gray Cary Ware & Freidenrich LLP  
; STREET: 4365 Executive Drive, Suite 1600  
; CITY: San Diego  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92121-2189  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/880,708  
; FILING DATE: 12-Jun-2001  
; PRIOR APPLICATION DATA:

;  
; APPLICATION NUMBER: 09/145,060  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: 08/003,144  
; FILING DATE: 12-JAN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lisa A. Haile, Ph.D.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 07265/057002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 858/677-1456  
; TELEFAX: 619/677-1465  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 495 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-09-880-708-10

Query Match 37.6%; Score 919.5; DB 9; Length 495;  
Best Local Similarity 43.5%; Pred. No. 4.6e-58;  
Matches 209; Conservative 74; Mismatches 107; Indels 91; Gaps 18;  
QY 21 PGFQQAISISSCSAELGSGTKMRSRKEGKQAPRDSADREGQGPQPPQDE----- 74  
Db 60 PGHYGVGATNARAK-GSSGQTQAKKDEPKMPPRSG-----GSETKGCPSSQTQAAA 113  
QY 75 -----PRAQEP-----RAQEP-----PGRGRVVPVPEHYMLSIYRT 104  
Db 114 RTWTPKQQLPGGKASSKAGSAPSSFLKKTRTREPPTREPKEFPPTPTPHEYMLSLYRT 173  
QY 105 YSIAEKLGINASFFQSSKSANTITSFVDRGLDDLSHTPLRQKYLFDVSMDSKEELVGA 164  
Db 174 LSDADRKGNSSVKLEAGLANITISFIDKGQDDRG--PAVRKQRYVFDISAL--EKDGLLGA 231  
QY 165 ELRLFRQAP--SAPMGPPAGPL-HVQLPCLLS---PLLLDARTLDPGAPAGWEVD 216  
Db 232 ELRLRKKPLDVAKPAVPSSGVRVAQLKLSSCSGQAPSLLDVRSV--PGLDGSQGEVD 289  
QY 217 VMOGLRH-QPWKQLCLLELRAANGELDAGAEARAGPQPPPPDLKSLGFGRRVRPQER 275  
Db 290 IWKLFNFKNLSAQLCLELE-AW---ERGRA-----VDLRLGLGFGERTARQVHEK 333  
QY 276 ALLVVFTRSORKNL-FAEMREQLGSAAAGPAGAGEGSWPPSPGADARPWLPSGRRRR 334  
Db 334 ALFLVFGRTKKRDLFFNEIKARSGQDDKT-----VVEYLFQ--RRKR 374  
QY 335 RTAFASRHGKRHGRKSRKPLHVNFKELGWDWIITAPLEYEAYHCEGVCDPFLRSH 394  
Db 375 RAPLANRQGRKPSKKNLAKRCSRKALHVNFKDGMWDWIITAPLEYEAFHCEGLCEFFLRSH 434  
QY 395 LEPTNHAIIQTLMNSMDPSTPTCCVPTKLPISILYIDAGNNVYKYEDMVVESCGR 454  
Db 435 LEPTNHAIIQTLMNSMDPESTPTCCVPTRLSPISILFIDSANNVYKYEDMVVESCGR 494  
QY 455 R 455  
Db 495 R 495

RESULT 7  
US-09-813-398-37  
; Sequence 37, Application US/09813398  
; Patent No. US20020169292A1  
; GENERAL INFORMATION:  
; APPLICANT: Bruce D. Weintraub  
; APPLICANT: Mariusz W. Szkudlinski  
; APPLICANT: University of Maryland  
; TITLE OF INVENTION: CYSTINE KNOT GROWTH FACTOR MUTANTS  
; FILE REFERENCE: USFMD.003C1  
; CURRENT APPLICATION NUMBER: US/09/813,398

```

; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSQL for Windows Version 2.0b
;

```

ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 620 Newport Center Drive, 16th Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
FILING DATE: US/09/735,849  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/836,081  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Bartfield, Neil S  
REGISTRATION NUMBER: 39,901  
REFERENCE/DOCKET NUMBER: NIH099.001APC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176  
TELEX:  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 501 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-735-849-13

Query Match 36.9%; Score 903.5; DB 10; Length 501;  
Best Local Similarity 44.9%; Pred. No. 6.3e-57;  
Matches 207; Conservative 63; Mismatches 108; Indels 83; Gaps 16;  
QY 38 GSTGMRSRKEGKMORAP--RSDAGREGQEPQPRQD---EPRAQQP----- 80  
Db 81 GQTGGLTQPKDEPKLPDPGPGEPKPPQTRQATRTVTPKQLPGGKAPPKAGSV 140  
QY 81 -----RAQEP-----PGRGPRVVPHEYMLSYRTYSIAEKLGINASFQSKSAN 125  
Db 141 PSSFLKKAREPGPREPPEPRPPITTHEYMLSIRTLSDADRKGGNSVKLEAGLAN 200  
QY 126 TTTSFVDRGLDLSHTPLRRQYLFVDSMLSKEELVGAELRLFRQAPSAPWGPAGP-- 183  
Db 201 TTTSFDKGQDGRGPV-VRKQRYVFDISAL-EKDGLLGAELRLRKPS-DTAKPAVPRS 257  
QY 184 ---LHVQLFPCLs---PLLLDARTLDPOGAPPAGWEVDFVMOGLRH-QPWKOLCLELRA 235  
Db 258 RRAAQKLLSCSPSGQPAALLDVRSV--PGLDGSQGEVDFIWKLFNFKNSAQLCLEL-- 313  
QY 236 AMGELDAGAEARAGPQPPPPDLRLSLGFRRRVPPQERALLVVFTRSORKNL-FAEMR 294  
Db 314 -----EAWERGR-----TVDLRGLGFDRAARQVHEKALFLVFGRTKKRDLFENEIK 359  
QY 295 EQLGSAEAGPAGAGSGSWPPSPGADARPWLPSPGRRRRRTAFASRRHGKRSRLRC 354  
Db 360 ARSGQDDKT-----VVEYLFQ--RRKRRAPSATROQKRPSKNLKARC 400  
QY 355 SKKPLHVNFKELGDDWIIAPLEYEAYHCEGVCDPPLRSHLEPTNHAIQTLLMSMDPGS 414  
Db 401 SKKALHVNFKMGDDWIIAPLEYEAFQGLCEPPLRSHLEPTNHAVIQTLLMSMDPS 460  
QY 415 TTPSCVPTKLPISILYIDAGNNVYKQYEDMVVESCGR 455  
Db 461 TPTCCVPTRLSPISILFIDSANNVYKQYEDMVVESCGR 501

RESULT 10

US-09-945-182-34  
Sequence 34, Application US/09945182  
Patent No. US2002016049A1  
GENERAL INFORMATION:  
APPLICANT: Celeste, Anthony J.  
Wozney, John  
Rosen, Vicki A.  
Wolfman, Neil  
Thomsen, Gerald H.  
Melton, Douglas A.  
TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENETICS INSTITUTE, INC.  
STREET: 87 CambridgePark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/945,182  
FILING DATE: 31-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION NUMBER:  
APPLICATION NUMBER: 08/808,324  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Lazar, Steven R.  
REGISTRATION NUMBER: 32,618  
REFERENCE/DOCKET NUMBER: 5202-D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617 498-8260  
TELEFAX: 617 876-5851  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 388 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 34:  
US-09-945-182-34

Query Match 31.3%; Score 765; DB 9; Length 388;  
Best Local Similarity 51.2%; Pred. No. 3e-47;  
Matches 172; Conservative 36; Mismatches 78; Indels 50; Gaps 10;  
QY 146 QKYLEFVDSMLSKEELVGAELRLFRQ-APSAPWGPAGPLHVQLFPCL-----LSPILLDAR 200  
Db 77 QSFLEFVDSLSNDADEVGAELRVLRGSPESGPGSWTSPPLLLLLSTCPGAARAPRLLYSR 136  
QY 201 TLDPOCAPAGWEVDFVMOGLRH-----OPWKOLCLELRAAMGELDAGAEARAGPQOP 255  
Db 137 AAEP--LVQQRWEAFVADAMRRHRREPFRPRAFCULLLRVAVGPV-----P 180  
QY 256 PPPDLRLSLGF---GRRVRPPQERALLVVFTRSORK-NLFAEMREQ---LGSAAEAGPAG 308  
Db 181 SPLALRRLGFGWGGGSAEERAVLVVSSRTQRKESLFRREIRAQAARALGAALASEP--- 237  
QY 309 AEGSWPPSPGADARPWLPSPGRRRRRTAFASRH-----GKRHGKSLRCSKKPL 359  
Db 238 -----LPDPGTGTASPRAVIGGRRRRRTALAGTRTAQSGGGGAGRGHGRGSRCSKKPL 292  
QY 360 HVNFKELGDDWIIAPLEYEAYHCEGVCDPPLRSHLEPTNHAIQTLLMSMDPGSTPPSC 419  
Db 293 HVDFKELGDDWIIAPLEYEAYHCEGLCDFPLRSHLEPTNHAIQTLLNSMAPDAPASC 352  
QY 420 CVPTKLTPLISILYIDAGNNVYKQYEDMVVESCGR 455  
||| :|:||||||| ||||||||| ||||||||| |||

Db 251 pDAAPASCCVPARI.SPTSTI.YTDAANNVVYKQYEDMVFEACGCR 294

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

	Query Match	25.6%	Score	627;	DB	9;	Length	411;
	Best Local Similarity	37.1%;	Pred.	No. 1.9e-37;				
	Matches	164;	Conservative	54;	Mismatches	136;	Indels	88;
	Gaps	17;						
QY	49 GKWORAPRDSADGREGQPQRPRDEPRAQQP-----RAQEPGRGPVVPHVWLS	100	:	:	:	:	:	:
Dd	23 GLVPGLRRKFAAASSGRPSSQPSDEVSLFEFLRLSMFGLKQBTPTSRDVAWP-PYMLD	81	:	:	:	:	:	:
QY	101 IYRTYSTAEKLGINAS-----FFQSXSANTITSF-VDRGLDLSHTPLLR-ROKYLFDV	152	:	:	:	:	:	:
Dd	82 LYRHRS-----GQPGSPADPHRLERAASTRNTVSFHHEESLEELPETSGKTTRRFENL	136	:	:	:	:	:	:
QY	153 SMLSCKEELVGAEURLFRQAPSAPGWGGAGPLH-VQLFPCLSPI-----LLDART	201	:	:	:	:	:	:
Dd	137 SSPTTEEFITSAEQLVFREQMDALGNSSFHHRIINTIYEIIKPATANSKFPTVRLDLLTRL	196	:	:	:	:	:	:
QY	202 LDPQGAPPAGWEVDVWGQRLRHOPWKOLCLELAAWGLDAGEAEAARFGPOPPDLR	261	:	:	:	:	:	:
Dd	197 VN-QNA--SRWSESDVTTPAVNR-W-----TAQGHANHGFTVVVEVAH-LEEKGQGVSKR	242	:	:	:	:	:	:
QY	262 SLGCFGRRVRP-----POERALLVVFTTRSQRKNLF--AEMREQL-GSAEAAGPGAGAEGSW	313	:	:	:	:	:	:

Db 243 HVRISRLHQDEHSWSQIRPLLVTFGHDKGHPHLHREKRTALAGTRTAQGGGGA----- 298

QY 314 PPSGAPDARPWLPSPGRRRRRTAFASRHGKRGKSLRCSKKPLHVNFKELGWDWII 373

Db 299 -----GRHGRGRGRCSKRPJHVDKELGWDWII 329

QY 374 APLEYAYHCEGVCDPPLRSHLEPTNHAIIQTLNMSDPGTPPSCCVPTKLTPTISILYI 433

Db 330 APLDYAYHCEGLCDPPLRSHLEPTNHAIIQTLNMSDPAAPASCCVPARLSPTISILYI 389

QY 434 DAGNNVYKQYEDMVVESCGR 455

Db 390 DAANNVYKQYEDMVVYACGR 411

RESULT 13

US-09-945-182-30

; Sequence 30, Application US/09945182

; Patent No. US20020160494A1

; GENERAL INFORMATION:

; APPLICANT: Celeste, Anthony J.

; Wozney, John

; Rosen, Vicki A.

; Wolfman, Neil

; Thomsen, Gerald H.

; Melton, Douglas A.

; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS

; NUMBER OF SEQUENCES: 35

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENETICS INSTITUTE, INC.

; STREET: 87 Cambridgepark Drive

; CITY: Cambridge

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/945,182

FILING DATE: 31-Aug-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/808,324

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Lazar, Steven R.

REGISTRATION NUMBER: 32,618

REFERENCE/DOCKET NUMBER: 5202-D

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617 498-8260

TELEFAX: 617 876-3851

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 240 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 30:

US-09-945-182-30

Query Match 25.3%; Score 618; DB 9; Length 240;

Best Local Similarity 55.6%; Pred.No.4.5e-37;

Matches 130; Conservative 20; Mismatches 34; Indels 50; Gaps 7;

QY 260 LRLSGFGRVRPP-----QERALLVFTRSQK-NLFAEMREQLSAAAGPGAG 308

Db 19 LRLGFGX----PGGDGGGTAXEERALLVISRTQKESLFEIRAQARLAA----- 69

QY 309 AEGSWPPSPGADPARWLPSPGRRRRRTAFASRH----- 342

Db 70 AE----PPDPGPGAGSRKANLGGRRRQRTALAGTRGXGSGGGGGGGGGGGGGG 126

QY 343 -GRRHGKSLRCSKKPLHVNFKELGWDWIIAPLEYAYHCEGVCDPPLRSHLEPTNHA 401

Db 127 AGRHGRGRGRCSKRLSHVDKELGWDWIIAPLDYAYHCEGVCDPPLRSHLEPTNHA 186

QY 402 IIQTLNMSDPGTPPSCCVPTKLTPTISILYIDAGNNVYKQYEDMVVESCGR 455

Db 187 IIQTLNMSDPAAPASCCVPARLSPTISILYIDANNVYKQYEDMVVYACGR 240

RESULT 14

US-08-945-459A-1

; Sequence 1, Application US/08945459A

; Patent No. US20020102633A1

; GENERAL INFORMATION:

; APPLICANT: MAKISHIMA, FUSAO; TAKAMATSU,

; APPLICANT: HIRIYUKI, MIKI, HIDEO; KAWAI,

; APPLICANT: SHINJI, KIMURA, MICHIO; MATSUMOTO,

; APPLICANT: TOMOAKI; KATSUURA, MIEKO; ENOMOTO,

; APPLICANT: KOICHI; SATOH, YUSUKE

; TITLE OF INVENTION: A NOVEL PROTEIN AND

; TITLE OF INVENTION: PROCESS FOR PREPARING THE SAME

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BIERMAN, MUSERLIAN AND LUCAS

; ADDRESSEE: LLP

; STREET: 600 THIRD AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10016

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: MICROSOFT WORD 97

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/945,459A

FILING DATE: 09-DEC-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP96/01062

FILING DATE: 19-APR-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP7/322403

FILING DATE: 17-NOV-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP7/93664

FILING DATE: 19-APR-1995

ATTORNEY/AGENT INFORMATION:

NAME: CHARLES A. MUSERLIAN

REGISTRATION NUMBER: 19,683

REFERENCE/DOCKET NUMBER: 146.1275

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 661-8000

TELEFAX: (212) 661-8002

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 119 AMINO ACIDS

TYPE: AMINO ACID

STRANDEDNESS:

TOPOLOGY: LINEAR

MOLECULE TYPE: PEPTIDE

ORIGINAL SOURCE:

ORGANISM: HOMOSAPIENS

TISSUE TYPE: FETUS

FEATURE:

NAME/KEY: MP52

LOCATION: 383 TO 501

US-08-945-459A-1

Query Match 22.4%; Score 549; DB 8; Length 119;







GenCore version 5.1.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2002, 02:50:50 ; Search time 91 Seconds  
(without alignments)  
1030.236 Million cell updates/sec

Title: US-09-825-751A-20

Perfect score: 2447

Sequence: 1 MDTPRVLLSAVFLISFLMDL.....GNVVYKQYEDMVVESCGR 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1253	51.2	399	13 Q9W753	Q9W753 xenopus lae
2	1091.5	44.6	412	13 Q12938	O12938 brachydanio
3	917	37.5	413	13 Q9DGN4	Q9DGN4 xenopus lae
4	902	36.9	500	13 Q9W6G0	Q9W6G0 gallus gall
5	870	35.6	447	6 Q9BDW8	Q9BDW8 cercopithec
6	841.5	34.4	324	13 Q9YHW9	Q9YHW9 gallus gall
7	737	30.1	294	6 Q9BDW9	Q9BDW9 macaca fasc
8	705	28.8	441	11 Q99MY1	Q99MY1 mus musculus
9	698.5	28.5	261	13 Q9W6C0	Q9W6C0 brachydanio
10	664.5	27.2	257	13 Q42303	O42303 brachydanio
11	612.5	25.0	126	13 Q93573	Q93573 gallus gall
12	507	20.7	417	5 Q9XVQ7	Q9XVQ7 lytechinus
13	457.5	18.7	391	13 Q8UVQ2	Q8UVQ2 brachydanio
14	456.5	18.7	391	13 Q8UVQ8	Q8UVQ8 brachydanio
15	453	18.5	398	13 Q90YD7	Q90YD7 xenopus tro
16	451	18.4	277	13 Q90Y82	Q90Y82 lampetra ja

17	448.5	18.3	391	13 Q90YJ3	Q90YJ3 brachydanio
18	445	18.2	400	13 Q57574	O57574 brachydanio
19	442	18.1	364	13 Q9PVK1	Q9PVK1 gallus gall
20	439.5	18.0	405	5 Q9U5E8	Q9U5E8 ptychodera
21	438	17.9	361	5 Q96504	Q96504 branchiosteo
22	438	17.9	400	13 Q13107	O13107 brachydanio
23	438	17.9	411	5 Q9U418	Q9U418 branchiosteo
24	431.5	17.6	390	13 Q91597	Q91597 xenopus lae
25	428.5	17.5	509	5 Q8WS99	Q8WS99 archaster t
26	426	17.4	289	5 Q9XVQ8	Q9XVQ8 strongyloce
27	424.5	17.3	386	13 Q13109	O13109 brachydanio
28	424	17.3	337	6 Q9MZV5	Q9MZV5 canis famil
29	412	16.8	614	5 Q91720	P91720 drosophila
30	409.5	16.7	400	13 Q73818	Q73818 xenopus lae
31	409.5	16.7	411	13 Q93369	Q93369 brachydanio
32	408.5	16.7	400	13 Q91703	Q91703 xenopus lae
33	405.5	16.6	453	13 Q87373	P87373 gallus gall
34	404.5	16.5	400	13 Q90YD6	Q90YD6 xenopus tro
35	403.5	16.5	411	13 Q57573	O57573 brachydanio
36	402	16.4	301	5 Q97390	Q97390 crassostrea
37	401.5	16.4	204	5 Q9XZ69	Q9XZ69 tripneustes
38	398	16.3	391	13 Q87380	P87380 brachydanio
39	397.5	16.2	432	13 Q9PTP9	Q9PTP9 brachydanio
40	396.5	16.2	398	13 Q918T6	Q918T6 gallus gall
41	395.5	16.2	588	5 Q9VQC6	Q9VQC6 drosophila
42	391	16.0	313	13 Q91403	Q91403 gallus gall
43	390.5	16.0	128	5 Q95W38	Q95W38 schistocerc
44	390	15.9	443	5 Q76851	Q76851 halocynthia
45	385.5	15.8	182	13 Q90Y81	Q90Y81 lampetra ja

#### ALIGNMENTS

RESULT 1

Q9W753 Q9W753 PRELIMINARY; PRT; 399 AA.  
 ID AC Q9W753;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Growth and differentiation factor 6.  
 GN GDF6.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99396700; PubMed=10393114;  
 RA Chang C., Hemmati-Briylanlou A.;  
 RT "Xenopus GDF6, a new antagonist of noggin and a partner of BMPs."  
 RL Development 126:3347-3357(1999).  
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 DR EMBL; AF155125; RAD38402.1; -.  
 DR HSSP; PI2643; 3BMP.  
 DR InterPro; IPR002405; Inhibin\_alpha.  
 DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFb\_propeptide; 1.  
 DR PRINTS; PR00669; INHIBINA.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFb; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 KW Glycoprotein.  
 SQ SEQUENCE 399 AA; 45571 MW; C549D973B50B8517 CRC64;

Query Match 51.2%; Score 1253; DB 13; Length 399;

Best Local Similarity 58.8%; Pred. No. 1.7e-92;

Matches 270; Conservative 40; Mismatches 85; Indels 64; Gaps 9;









```
Db 58 VSDSNNEAVHFGM-----LGLSREDQOQTHERRALLVAFSQARRKENLFREIREKIRAM 109
Qy 301 EAAGPGAGAGSWPPSPGAPDAPWLPSPGRRRTAFASR-----HGKRGHKKSLR 353
Db 110 KS-----RKFSNPTPEHSIKGHP-----RHRRTTALAGRPGVGPITSGGGGRRT 159
Qy 354 CSKPLHVNFKELGWDWIIAPLEYAYHCEGVCDFPLRSHLEPTNHAIQTOLMNSMDPG 413
Db 160 CSKPLHVNFKELGWDWIIAPLDYAYHCEGLCDFPLRSHLEPTNHAIQTOLMNSMDPE 219
Qy 414 STPPSCCVPTKLTPTISILYIDAGNNVVKQYEDMVVESCGR 455
Db 220 STPPSCCVPSKLSILYIDSGNNVVKQYEDMVVESCGR 261

RESULT 10
Q42303 ID 042303 PRELIMINARY; PRT; 257 AA.
AC 042303;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DE Contact (Fragment).
GN GDF5.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97398455; PubMed=9256353;
RA Bruneau S., Mourrain P., Rosa F.M.;
RT "Expression of contact, a new zebrafish DVR member, marks mesenchymal
RT cell lineages in the developing pectoral fins and head and is
RT regulated by retinoic acid.";
RL Mech. Dev. 65:163-173(1997).
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR EMBL; Y12005; CAA72733.1; -.
DR HSSP; P12643; 3BMP.
DR ZFIN; ZDB-GENE-990415-39; gdf5.
DR InterPro; IPR002400; GF_cysknot.
DR Pfam; PF00019; TGF-beta; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR PRODOM; PD000357; TGFb; 1.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Glycoprotein.
FT NON_TER 1
FT CHAIN 140 257
FT SEQUENCE 257 AA; 29787 MW; 6D64F0542F948849 CRC64;

Query Match 27.2%; Score 664.5; DB 13; Length 257;
Best Local Similarity 51.8%; Pred. No. 1.6e-45;
Matches 144; Conservative 32; Mismatches 53; Indels 49; Gaps 9;

Qy 186 VOLFPCLSL-----PLLDARTLDPOGAPGAGVEFDVW---QGLRHQPKQLCLELRAAG 238
Db 21 LRLFTCASGKNAVLQARPDFSHA--SYVEVFDIWKVFNFRNTP--QJCLELDA--- 73

Qy 239 ELDAEAEARAGPOQPPPLDLRSFGRRVRPGERALLVVFTRSORKNLF-ADMRRQL 297
Db 74 -VDHGR-----PLDLRLGLSRAGQTKEAFFVVGFTKKRGLFYNEIKAR- 119

Qy 298 GSAEAGPGAGAGSWPPSPGAPDAPWLPSPGRRRTAFASRHHGKHSRLRCSKK 357
Db 120 -----SCHDNKTVEYLFQTRMRRAPLPR-GKPKIPKQRCNRK 159

Qy 358 PLHVNFKELGWDWIIAPLEYAYHCEGVCDFPLRSHLEPTNHAIQTOLMNSMDPGSTPP 417
Db 160 QHVNFKELGWDWIIAPLEYAFHCDGVCDFPIKSHLEPTNHAIQTOLMNSMDPRSTPP 219
```

```
Qy 418 SCCVPTKLTPTISILYIDAGNNVVKQYEDMVVESCGR 455
Db 220 TCCVPTLSPISILYIDSGNNVVKQYEDMVVESCGR 257

RESULT 11
Q93573 ID 093573 PRELIMINARY; PRT; 126 AA.
AC 093573;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Putative growth/differentiation factor 6/7 (Fragment).
GN GDF6/7.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99026113; PubMed=9808626;
RA Lee K.J., Mendelsohn M., Jessell T.M.;
RT "Neuronal patterning by BMPs: a requirement for GDF7 in the generation
RT of a discrete class of commissural interneurons in the mouse spinal
RT cord.";
RL Genes Dev. 12:3394-3407(1998).
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR EMBL; AF089086; AAC97113.1; -.
DR HSSP; P12643; 3BMP.
DR InterPro; IPR002400; GF_cysknot.
DR InterPro; IPR001839; TGFb.
DR Pfam; PF00019; TGF-beta; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR PRODOM; PD000357; TGFb; 1.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Glycoprotein.
FT NON_TER 1
FT SEQUENCE 126 AA; 14265 MW; CB824D280F44A394 CRC64;

Query Match 25.0%; Score 612.5; DB 13; Length 126;
Best Local Similarity 86.5%; Pred. No. 1e-41;
Matches 109; Conservative 10; Mismatches 6; Indels 1; Gaps 1;

Qy 331 RRRRTAFASRHGKR-HGKHSRLRCSKKPLHVNFKELGWDWIIAPLEYAYHCEGVCDF 389
Db 1 RRRRTTIAARSGRGHGKAKTRCSRKPLHVNFKELGWDWIIAPLDYAYHCEGVCDF 60

Qy 390 PLRSHLEPTNHAIQTOLMNSMDPGSTPPSCCVPTKLTPTISILYIDAGNNVVKQYEDMVV 449
Db 61 PLRSHLEPTNHAIQTOLMNSMDPESTPPSCCVPSKLSILYIDSGNNVVKQYEDMVV 120

Qy 450 ESCGCR 455
Db 121 ETCGCR 126

RESULT 12
Q9XYQ7 ID 09XYQ7 PRELIMINARY; PRT; 417 AA.
AC 09XYQ7;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Bone morphogenetic protein BMP2/4.
GN BMP2/4.
OS Lytechinus variegatus (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidae; Euechinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;
OC Lytechinus.
```







GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2002, 02:33:40 ; Search time 37 seconds  
(without alignments)  
510.047 Million cell updates/sec

Title: US-09-825-751A-20

Perfect score: 2447

Sequence: 1 MDTPRVLLSAVFLISFLMDL.....GNVVYKQYEDMVVESQGR 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query %	Match	Length	DB ID	Description
1	1798	73.5	436	1	GDF6_BOVIN	P55106 bos taurus
2	927.5	37.9	501	1	GDF5_HUMAN	P43026 homo sapien
3	918.5	37.6	495	1	GDF5_MOUSE	P43027 mus musculus
4	691	28.2	125	1	GDF6_MOUSE	P43028 mus musculus
5	549	22.4	151	1	GDF7_MOUSE	P43029 mus musculus
6	471	19.2	398	1	BMP4_XENLA	P25703 xenopus lae
7	464	19.0	398	1	BMP8_XENLA	P30884 xenopus lae
8	453	18.5	405	1	BMP4_CHICK	Q90752 gallus gall
9	449.5	18.4	395	1	BMP2_RABIT	O46564 oryctolagus
10	441.5	18.0	408	1	BMP4_HUMAN	P12644 homo sapien
11	436.5	17.8	408	1	BMP4_MOUSE	P21275 mus musculus
12	434	17.7	393	1	BMP2_RAT	P49001 rattus norv
13	432.5	17.7	408	1	BMP4_RAT	Q06826 rattus norv
14	431.5	17.6	396	1	BMP2_HUMAN	P12643 homo sapien
15	430	17.6	409	1	BMP4_RABIT	O46576 oryctolagus
16	429.5	17.6	461	1	DVR1_STRPU	P48969 strongyloce
17	427.5	17.5	396	1	BMP2_DAMDA	O19006 dama dama
18	423	17.3	394	1	BMP2_MOUSE	P21274 mus musculus
19	419.5	17.1	408	1	BMP4_DAMDA	Q25607 dama dama
20	418.5	17.1	452	1	BMP5_MOUSE	P49003 mus musculus
21	418	17.1	424	1	BM10_HUMAN	O95393 homo sapien
22	417.5	17.0	353	1	BMP2_CHICK	Q90751 gallus gall
23	416.5	17.0	402	1	BMP8_HUMAN	P34820 homo sapien
24	416.5	17.0	454	1	BMP5_HUMAN	P22003 homo sapien
25	413.5	16.9	401	1	BMP4_XENLA	P30885 xenopus lae
26	412.5	16.9	399	1	BM8A_MOUSE	P34821 mus musculus
27	407	16.6	426	1	BMP7_XENLA	P30886 xenopus lae
28	406	16.6	420	1	BM10_MOUSE	Q91229 mus musculus
29	403	16.5	621	1	DECA_DROPS	P91699 drosophila
30	401.5	16.4	431	1	BMP7_HUMAN	P18075 homo sapien
31	400	16.3	430	1	BMP7_MOUSE	P23359 mus musculus
32	398	16.3	428	1	GDF2_MOUSE	Q9WV56 mus musculus
33	395.5	16.2	588	1	DECA_DROME	P07713 drosophila

#### RESULT 1

ID	GDF6_BOVIN	STANDARD;	PRT;	436 AA.
AC	P55106;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Growth/differentiation factor 6 precursor (GDF-6) (Cartilage-derived morphogenetic protein 2) (CDMP-2) (Fragment).			
GN	GDF6 OR CDMP2.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Articular cartilage;			
RX	MEDLINE=95050604; PubMed=7961761;			
RA	Chang S., Hoang B., Thomas J.T., Vukicevic S., Luyten F.P.,			
RA	Ryba N.J.P., Kozak C.A., Reddi A.H., Moos M.;			
RT	"Cartilage-derived morphogenetic proteins. New members of the transforming growth factor-beta superfamily predominantly expressed in long bones during human embryonic development.";			
RL	J. Biol. Chem. 269:28227-28234(1994).			
CC	-!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).			
CC	-!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	EMBL; U13661; AAA61416.1; -			
DR	HSSP; P18075; IEMP.			
DR	InterPro; IPR001839; TGFb.			
DR	InterPro; IPR001111; TGFb.N.			
DR	Pfam; PF00019; TGF-beta.1			
DR	Pfam; PF00688; TGFb.propeptide; 1.			
DR	ProDom; PD000357; TGFb; 1.			
DR	SMART; SM00204; TGFb; 1.			
DR	PROSITE; PS00250; TGF-BETA.1; 1.			
KW	Growth factor; Cytokine; Glycoprotein.			
FT	NON_TER	1		
FT	PROPEP	<1	316	POTENTIAL.
FT	CHAIN	317	436	GROWTH/DIFFERENTIATION FACTOR 6.
FT	DISULFID	335	401	BY SIMILARITY.
FT	DISULFID	364	433	BY SIMILARITY.
FT	DISULFID	368	435	BY SIMILARITY.
FT	DISULFID	400	400	INTERCHAIN (BY SIMILARITY).
FT	CARBOHYD	27	27	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	89	89	N-LINKED (GLCNAC. .) (POTENTIAL).
SQ	SEQUENCE	436 AA;	47873 MW;	B0688E12EF8AE91D CRC64;

#### ALIGNMENTS

```
Query Match          73.5%; Score 1798; DB 1; Length 436;
Best Local Similarity 80.8%; Pred. No. 4.2e-113;
Matches 354; Conservative 18; Mismatches 48; Indels 18; Gaps 8;

QY 33 SBAELGSKGMSRKEGKQAPRSDAGRE--QEQPQPRQDDEPR---AQOPRAQEP 86
DB 2 ASAEIGSAGKMRKEGRMPRAPRENATAREPLDRQEPQPRQEPQPRQEPQEP 61

QY 87 GGRPVVPEHYMLSTYRTYSIAEKLGINASFQSSKSANTITSFVDRGLDDLSHTPLRQ 146
DB 62 GGRPLVPPEHYMLSTYRTYSIAEKLGINASFQSSKSANTITSFVDRGLDDLSHTPLRQ 121

QY 147 KYLFVDSMLSDKEELVGAELRFRQAPSPWGPAGPLHVQLFPCPLSLLLDARTLDPOG 206
DB 122 KYLFVDSMLSDKEELVGAELRFRQAPSPWGPAGPLHVQLFPCPLSLLLDARTLDPOG 179

QY 207 APAGWEVDFVQGLRHQPKWQLCLLELRAW--GELDAGEAERARQPPQPPDRLSLGF 265
DB 180 APRPGWEVDFVQGLRHQPKWQLCLLELRAWGPGEAERARQPPQPPDRLSLGF 239

QY 266 GRRVRPQERALLVFTSRQNLFAEMREQLGSA--EAGAGCGAGAGSW-----PPPS 317
DB 240 GRRVRPQERALLVFTSRQNLFAEMREQLGSA--EAGAGCGAGAGSW-----PPPS 299

QY 318 GAPDARPNLPSGRRRRRTAFASRHGKRGKSRKRLRCSKKPLHVNFKELGWDWIIAPLE 377
DB 300 GTPDAGLWSPSG-RRRTAFASRHGKRGKSRKRLRCSKKPLHVNFKELGWDWIIAPLE 358

QY 378 YEAYHCEGVCDFPLRSHLEPHTHAIQIOLMSMDGSPSCCVPTKLTPIILYIDAGN 437
DB 359 YEAYHCEGVCDFPLRSHLEPHTHAIQIOLMSMDGSPSCCVPTKLTPIILYIDAGN 418

QY 438 NVVYKQYEDMVYESGCR 455
DB 419 NVVYKQYEDMVYESGCR 436

RESULT 2
GDF5_HUMAN
ID GDF5_HUMAN STANDARD; PRT; 501 AA.
AC P43026; Q96SBI;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Growth/differentiation factor 5 precursor (GDF-5) (Cartilage-derived
DE morphogenetic protein 1) (CDMP-1).
DE GDF5 OR CDMP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=95071375; PubMed=7980526;
RA Hoetten G., Neidhardt H., Jacobowsky B., Pohl J.;
RT "Cloning and expression of recombinant human growth/differentiation
RT factor 5.";
RL Biochem. Biophys. Res. Commun. 204:646-652(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Articular cartilage;
RX MEDLINE=95050604; PubMed=7961761;
RA Chang S., Hoang B., Thomas J.T., Vukicevic S., Luyten F.P.,
RA Ryba N.J.P., Kozak C.A., Reddi A.H., Moos M.;
RT "Cartilage-derived morphogenetic proteins. New members of the
RT transforming growth factor-beta superfamily predominantly expressed
RT in long bones during human embryonic development.";
RL J. Biol. Chem. 269:28227-28234(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
```

```
FT CONFLICT 276 276 S -> A (IN REF. 2).
FT CONFLICT 321. 321 A -> T (IN REF. 2).
FT CONFLICT 384 384 L -> S (IN REF. 2).
SQ SEQUENCE 501 AA; 55410 MW; 37985F2D15C4F5EF CRC64;

Query Match 37.9%; Score 927.5; DB 1; Length 501;
Best Local Similarity 44.7%; Pred. No. 5.4e-55;
Matches 209; Conservative 64; Mismatches 98; Indels 97; Gaps 16;

QY 38 GSTKMRSGKRGKQAPRDSADAGREGQBPQPPQDPQBPQAPQ----- 79
Db 81 QGTGGLTPQKKDEPKLP-----PRG-GPEKPGHPQTRQATRTVTPKGLPGCKAP 134
QY 80 -----PRAQEPGGRVVPVHEYMISYIAEKLGINASFFQ 119
Db 135 PKAGSVPSFLLKKAREPGPREPKPEPPPTPHEYMISLYRTLSDADRGKGNSSVKL 194
QY 120 SSKSANTITSFVDRGLDLSHTPLRQKYLFDVMSLDEELVGAELRFRQAPS----- 174
Db 195 EAGLANTITSFIDKGQDDRGV-VRKQYVFDISAL-ERDGLLGAELRLRKKPSDTAKP 252
QY 175 -APWGPAGPLHVLPFCLS-----PLLLDARTLDPOGAPGAGWEVDVWQGLRH-QPWKQ 228
Db 253 AAPGGRRAQL--KLSCSGROPASLLDVRV--PGLDGSWEVDIWKLFNFKNASQ 308
QY 229 LCLELRAANGELDAGAEARAGPOPPPPDLRLSLGFGRRVPPQERALLVVFTRSQKN 288
Db 309 LCLELE-AW---ERGRA-----VLRGLGFDRAARQVHEKALFLVFGRTKRD 352
QY 289 L-FAENREQLGSNAAGPAGAGAGEGSWPPSPGAPDARWLPSPGRRRRRTAFASRHGRHG 347
Db 353 LFNELTKARSGQDDKT-----VVEYLFQ--RRRRRAPLATRQGRPS 393
QY 348 KKSRLCSKKPLHVNFKELGWDWIIAPLEYEAYHCEGVCDDPLRSHLEPTNHAIIQTLM 407
Db 394 KNLKARCSKALHVNFKDGMWDWIIAPLEYEAFHCEGLCEPLRSHLEPTNHAIIQTLM 453
QY 408 NSMDPGSTPPCCVPTKLPISILYIDAGNNVYKQYEDMVVESCGR 455
Db 454 NSMDPESTPTCCVPTRLSPISILFDSANNVYKQYEDMVVESCGR 501

RESULT 3
GDF5_MOUSE STANDARD; PRT; 495 AA.
AC P43027;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Growth/differentiation factor 5 precursor (GDF-5).
GN GDF5 OR GDF-5 OR BP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=CD-1; TISSUE=Embryo;
RX MEDLINE=94195427; PubMed=8145850;
RA Storm E.E., Huynh T.V., Copeland N.G., Jenkins N.A., Kingsley D.M.,
RA Lee S.-J.;
RT "Limb alterations in brachypodism mice due to mutations in a new
RL Nature 368:639-643(1994).
CC -!- FUNCTION: COULD BE INVOLVED IN BONE FORMATION.
CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC -!- DISEASE: DEFECTS IN GDF5 ARE THE CAUSE OF BRACHYPODISM WHICH
CC ALTERS THE LENGTH AND NUMBERS OF BONES IN THE LIMBS BUT SPARES THE
CC AXIAL SKELETON.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; 008337; AAA18778.1; -
DR HSSP; P12643; 3BMP.
DR MGD; MGI:95688; Gdf5.
DR InterPro; IPR002400; GF_cysknot.
DR InterPro; IPR001839; TGFb.
DR InterPro; IPR001111; TGFb_N.
DR Pfam; PF00019; TGF-beta; 1.
DR Pfam; PF00688; TGFb_propeptide; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR ProDom; PD000357; TGFb; 1.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Signal; Growth factor; Cytokine; Glycoprotein; Polymorphism.
FT SIGNAL 1 27 POTENTIAL.
FT PROPEP 28 375 POTENTIAL.
FT CHAIN 376 495 GROWTH/DIFFERENTIATION FACTOR 5.
FT DISULFID 394 460 BY SIMILARITY.
FT DISULFID 423 492 BY SIMILARITY.
FT DISULFID 427 494 BY SIMILARITY.
FT DISULFID 459 459 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 183 183 N-LINKED (GLCNAC..). (POTENTIAL).
FT VARIANT 98 98 S -> P.
SQ SEQUENCE 495 AA; 54885 MW; CD0D5DE48185D2E3 CRC64;

Query Match 37.6%; Score 919.5; DB 1; Length 495;
Best Local Similarity 43.5%; Pred. No. 1.8e-54;
Matches 209; Conservative 74; Mismatches 107; Indels 91; Gaps 18;

QY 21 PGFQASISSSCSAELGSKMRSRKEKMQAPRDSADAGREGQBPQPPQDE----- 74
Db 60 PGHIVGVGATNARAK-GSSGQTQAKDKPRKMPPRSG-----GSETKFGPSQTQAAA 113
QY 75 ----PRAQOP-----RAQEP-----PGRGPRVVPHEYMISYRT 104
Db 114 RVTPKQLPGGKASKKAGSAPSSFLKKTTREPTPREKPEPRPTTTPHETWLSYRT 173
QY 105 YSTAELKGINASFFQSSKSANTITSEVDRLDLSHTPLRRQKYLFDVMSLDEELVGA 164
Db 174 LSDADRGKGNSSVKLEAGLANITITSFIDKGQDDRG-PAVRKQRYVFDISAL-EKDGLGA 231
QY 165 ELRLFRQAP---SAPWGPAGPL-HVOLFPCLS-----PLLLDARTLDPOGAPGAGWEVD 216
Db 232 ELRLRKKPLDVAKPAVPSGSRVAQLKLSGSCPSGROFAALLDVRV--PGLDGSWEVD 289
QY 217 VWQGLRH-OPWKQLCLELRAANGELDAGAEARARGPQQPPPPDLRLSLGFGRRVPPQER 275
Db 290 INKLFNFNKSQLCLELE-AW---ERGRA-----VDLRGLGFERTAROVHEK 333
QY 276 ALLVVTFRSQKNL-FAENREQLGSAAAGAGAGSGSWPPSPGAPDARWLPSPGRRRR 334
Db 334 ALFLVFGRTKKRDLFFNEIKARSGQDDKT-----VVEYLFQ--RRR 374
QY 335 RTAFASRHGRKGRKSRRLCSKPLHVNFKELGWDWIIAPLEYEAYHCEGVCDFPLRSH 394
Db 375 RAPLANRQGRPSKNLKRCSKALHVNFKDGMWDWIIAPLEYEAFHCEGLCEFFLRSH 434
QY 395 LEPTNHAIQTLMNSMDPGSTPPCCVPTKLPISILYIDAGNNVYKQYEDMVVESCGR 454
Db 435 LEPTNHAIQTLMNSMDPESTPTCCVPTRLSPISILFDSANNVYKQYEDMVVESCGR 494
QY 455 R 455
Db 495 R 495

RESULT 4
GDF6_MOUSE
```

ID	GDF6_MOUSE	STANDARD;	PRT;	125 AA.
DT	P43028;			
AC	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Growth/differentiation factor 6 precursor (GDF-6) (Fragment).			
GN	GDF6 OR GDF-6.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BALB/C; TISSUE=Liver;			
RC	MEDLINE=941195427; PubMed=8145850;			
RA	Storm E.E., Huyh T.V., Copeland N.G., Jenkins N.A., Kingsley D.M.,			
RA	Lee S.-J.;			
RT	"Limb alterations in brachypodism mice due to mutations in a new			
RL	member of the TGF beta-superfamily.";			
RL	Nature 368:639-643(1994).			
CC	-1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).			
CC	-1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
EMBL	U08338; AAA18779.1; -			
DR	HSSP; P12643; 3BMP.			
DR	MGD; MGI:95689; Gdf6.			
DR	InterPro: IPR001839; TGFb.			
DR	Pfam: PF00019; TGF-beta; 1.			
DR	ProDom: PD000357; TGFb; 1.			
DR	SMART; SM00204; TGFb; 1.			
DR	PROSITE; PS00250; TGF_BETA_1; 1.			
KW	Growth factor; Cytokine; Glycoprotein.			
FT	NON_TER 1			
FT	PROPEP <1 5			
FT	CHAIN 6 125			
FT	DISULFID 24 90			
FT	DISULFID 53 122			
FT	DISULFID 57 124			
FT	DISULFID 89 89			
FT	DISULFID 89 89			
FT	INTERCHAIN (BY SIMILARITY).			
FT	SEQUENCE 125 AA; 14373 MW; 10FA2A5B7748DA32 CRC64;			
FT	-----			
Query Match	28.2%; Score 691; DB 1; Length 125;			
Best Local Similarity	99.2%; Pred. No. 7.7e-40;			
Matches 124; Conservative	1; Mismatches 0; Indels 0; Gaps 0;			
QY	331 RRRRTAFASRHKRGKSKRLRCSSKPLHVNFKELGWDWIITAPLEYAYHCEGVCDFF 390			
DB				
DB	1 RRRRTAFASRHKRGKSKRLRCSSKPLHVNFKELGWDWIITAPLEYAYHCEGVCDFF 60			
QY	391 LRSHLEPTNHAIIQTLMNSMDPGSTPPSCCVPKLTPISTILYIDAGNNVYKQYEDMVE 450			
DB				
DB	61 LRSHLEPTNHAIIQTLMNSMDPGSTPPSCCVPKLTPISTILYIDAGNNVYKQYEDMVE 120			
QY	451 SCGR 455			
DB				
DB	121 SCGR 125			
RESULT 5				
GDF7_MOUSE				
ID	GDF7_MOUSE	STANDARD;	PRT;	151 AA.
ID	P43029;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			

[illegible]







```
DR PIR; C37278; C37278.
DR HSP; P12643; 3BMP.
DR Genew; HGNC:1071; BMP4.
DR MIM; 112262; -.
DR InterPro; IPR001839; TGFb.
DR InterPro; IPR001111; TGFb_N.
DR Pfam; PF00019; TGF-beta; 1.
DR Pfam; PF00688; TGFb-propeptide; 1.
DR ProDom; PD000357; TGFb; 1.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; P500250; TGF_BETA_1; 1.
KW Signal; Growth factor; Cytokine; Bone; Cartilage; Glycoprotein.
FT SIGNAL 1 19
FT PROPEP 20 292
FT CHAIN 293 408
FT DISULFID 308 373
FT DISULFID 337 405
FT DISULFID 341 407
FT DISULFID 372 372
FT CARBOHYD 143 143
FT CARBOHYD 208 208
FT CARBOHYD 350 350
FT CARBOHYD 365 365
FT CONFLICT 152 152
SQ SEQUENCE 408 AA; 46555 MW; 79B01179DBB98204 CRC64;

Query Match 18.0%; Score 441.5; DB 1; Length 408;
Best Local Similarity 28.0%; Pred. No. 1.1e-22;
Matches 135; Conservative 70; Mismatches 164; Indels 113; Gaps 19:

QY 7 LLSAVLFLSLDLRGFQOASISSCSAELGSLTKG-MRSREKGMQRAPRDSDA--GR 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7 MLMVLLCQVL--LGASHASLIPETGKKVAEIQGHAGRGSGOSHELLRFEATLQM 64
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 63 EQGEQPPQDPEPRAQPPQRAQPPGPPGPPVPPHEMYLSIYRYSIAEK-----LGINAS 116
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 65 FGLRRPQPSKS-----AVIP-DYMRDLYRLQSGEEDEEIQHSTGLEYP 107
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 117 PFQSSKSANTITSF-VDRGLDLSHTPLRRQ-KYLFVDSMLSKDELGAELRLFRQAPS 174
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 108 ERPASR-ANTVRSFHEEHELENIPGTSNSAFRFLNLSIPENEVISAEALRLFRE--Q 164
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 175 APWGP--PAGPLVHVFPLCLSP-----LLDARTLDPOGAPPAGNVEFVDMOGL 221
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 165 VQQGPDPWGEGRFINIYEMKPPAEVVPGLHITRLDTRLVHNH--VTRWETFDV--- 217
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 222 RHQPKQLCLELRAWGLDAGEAEARGPQPPDLRLSLGFGRRVPPPOERALLVVF 281
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 218 -----SPAVLR---WTRKQPNYGLAIEVTH 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 282 TRSQKNLFAEMREQLGSAEAGAGAGSGWPPPSGAPDARPLPSPGR-----RR 333
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 LHQTRTHOGHVRSLSLPQGS-----GNW-----AQLRPLVTGHDGRHALTRR 287
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 334 RTAFASRHGKRGKSRKLRCSKPLHNFKELGWDWDWIAPLEYAYHCEGCVDFPLRS 393
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 288 RRAKSPKHQSQRARKKNCRRHSILYVDFSDVGNWDWIVAPPGVQAFYCHGDCFPPLAD 347
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 394 HLEPTNHAIQTLMNSMDPGSTPPSCCVPTKLTPTISILYIDAGNVVYQYEDMVVESC 453
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 348 HLNSTNHAIQTVLNSVN-SSIPKACCVPTLSLMLYLDYDKVVLKLNQEMVEGCG 406
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 454 CR 455
   : :
Db 407 CR 408

RESULT 11
BMP4_MOUSE
ID BMP4_MOUSE
AC P21275;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
```

```
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bone morphogenetic protein 4 precursor (BMP-4) (BMP-2B).
GN BMP4 OR BMP-4 OR DVR-4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Dickinson M.E., van der Meer-De Jong R., Hogan B.L.M.;
RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93282803; PubMed=8507180;
RA Kurihara T., Kitamura K., Takaoka K., Nakazato H.;
RT "Murine bone morphogenetic protein-4 gene: existence of multiple
promoters and exons for the 5'-untranslated region.";
RL Biochem. Biophys. Res. Commun. 192:1049-1056(1993).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93365172; PubMed=8358941;
RA Takaoka K., Yoshikawa H., Hasimoto J., Masuhara K., Miyamoto S.,
RA Suzuki S., Ono K., Matsui M., Oikawa S., Tsuruoka N.;
RT "Gene cloning and expression of a bone morphogenetic protein derived
from a murine osteosarcoma.";
RL Clin. Orthop. Relat. Res. 294:344-352(1993).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv; TISSUE=Liver;
RX MEDLINE=96081880; PubMed=7499338;
RA Feng J.Q., Chen D., Cooney A.J., Tsai M., Harris M.A., Tsai S.Y.,
RA Feng M., Mundy G.R., Harris S.E.;
RT "The mouse bone morphogenetic protein-4 gene. Analysis of promoter
utilization in fetal rat calvarial osteoblasts and regulation by
COUP-TF1 orphan receptor.";
RL J. Biol. Chem. 270:28364-28373(1995).
RN [5]
RP SEQUENCE OF 241-408 FROM N.A.
RX MEDLINE=90228966; PubMed=1970330;
RA Dickinson M.E., Kobrin M.S., Silan C.M., Kingsley D.M., Justice M.J.,
RA Miller D.A., Ceci J.D., Lock L.F., Lee A., Buchberg A.M.,
RA Siracusa L.D., Lyons K.M., Derynck R., Hogan B.L.M., Copeland N.G.,
RA Jenkins N.A.;
RT "Chromosomal localization of seven members of the murine TGF-beta
superfamily suggests close linkage to several morphogenetic mutant
loci".
RL Genomics 6:505-520(1990).
CC -!- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION.
CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED INTO THE EXTRACELLULAR MATRIX.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; X56848; CAA40179.1; -.
DR EMBL; S65032; AAB28021.1; -.
DR EMBL; L47480; AAC37698.1; ALT_INIT.
DR EMBL; D14814; BAA03555.1; -.
DR PIR; B34201; B34201.
DR HSSP; P12643; 3BMP.
DR MGD; MGI:88180; Bmp4.
DR InterPro; IPR001839; TGFb.
DR InterPro; IPR001111; TGFb_N.
DR Pfam; PF00019; TGF-beta; 1.
DR Pfam; PF00688; TGFb-propeptide; 1.
DR ProDom; PD000357; TGFb; 1.
```





BMP4\_RAT  
ID BMP4\_RAT STANDARD; PRT; 408 AA.  
AC Q06826;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Bone morphogenetic protein 4 precursor (BMP-4) (BMP-2b).  
GN BMP4 OR BMP-4 OR DVR-4.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93385158; PubMed=8373807;  
RA Chen D., Feng J.Q., Feng M., Harris M.A., Mundy G.R., Harris S.B.;  
RT "Cloning and sequence of bone morphogenetic protein 4 cDNA from fetal  
rat calvarial cell.";  
RL Biochim. Biophys. Acta 1174:289-292(1993).  
CC !- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION.  
CC !- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).  
CC !- SUBCELLULAR LOCATION: SECRETED INTO THE EXTRACELLULAR MATRIX.  
CC !- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: 222607; CAA80329.1; -  
DR PIR: S33173; S33173.  
DR PIR: S38343; S38343.  
DR HSP: P12643; 3BMP.  
DR InterPro: IPR001839; TGFb.  
DR InterPro: IPR001111; TGFb.N.  
DR Pfam: PF00019; TGF-beta; 1.  
DR Pfam: PF00688; TGFb.propeptide; 1.  
DR ProDom: PD000357; TGFb; 1.  
DR SMART: SM00204; TGFb; 1.  
DR PROSITE: PS00250; TGF\_BETA\_1; 1.  
KW Signal; Growth factor; Cytokine; Bone; Cartilage; Glycoprotein.  
FT SIGNAL 1 19 POTENTIAL.  
FT PROPEP 20 292  
FT CHAIN 293 408 BONE MORPHOGENETIC PROTEIN 4.  
FT DISULFID 308 373 BY SIMILARITY.  
FT DISULFID 337 405 BY SIMILARITY.  
FT DISULFID 341 407 BY SIMILARITY.  
FT DISULFID 372 407 INTERCHAIN (BY SIMILARITY).  
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 365 365 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 408 AA; 46540 MW; 61E92B4B8D5624F3 CRC64;  
Query Match 17.7%; Score 432.5; DB 1; Length 408;  
Best Local Similarity 29.1%; Pred. NO. 4.5e-22;  
Matches 138; Conservative 62; Mismatches 168; Indels 107; Gaps 19;  
QY 20 LPGFQQAQSSSCSAELGKMGSRKRGKQMRAPRDSAG--REGQEPQPRQDEPRA 77  
DB 2 IPGNRLAVLVLCVQLLGATDASLIPETGKKVAIEIQHAGGRSGSHELLRFEATL 61  
QY 78 QQP-----RAQEPGRGVVPVPEHYMLSYIYTVSIABK-----IGINASFQSSKS-ANTIT 128  
DB 62 LQMFGLRRRPPSKS-AVIP-DYMRDLVRLQSGEEEEEQSQGTGLEYPERSANTVR 119  
QY 129 SFVDRGLDLSHTPLRRQ-----KYLFDVYSMLSDKELVGAELRLFR----QAPSAPWG-- 178  
DB 120 SPFH--EHLNIPGTSBSSAFRRFFNLSSIPENVEISSAELRLRFRQVDGPDWEQGFH 177

QY 179 -----PPAG---PLHVQLFPCLSPLLLDARTLDPOGAPPAGWEVFDVWQGLRHQPWK 227  
DB 178 RNIYEVKPPAEVMPGH-----LITRLDLRLVRHN---VTRWETFV----- 218  
QY 228 QLCLELRAAWGELDAGEAEARAGPQQPPDLRSILGFGRRVRRPQERALLVVFTRSQRK 287  
DB 219 -----SPAVLR---WTREKQPNYGLAIEVTHLHQTRT 247  
QY 288 NLFAMREOLGSAEAGPGAGAGSGSWPPSPGAPDARPLPSPGR-----RRRTAFAS 340  
DB 248 HQGHVTRISRSLPQGS-----GNW-----AQLRPLLVTFEGHGRHTLRRRAKRS 294  
QY 341 RHGRHGRKSRCLKPLHVNFKELGWDDWITIAPLEYEAHCYGVCDPFLSHLEPTNH 400  
DB 295 KHPORSRKKKNCRHSLYVDFSDVGNWDWIVAPPGYQAFYCHGDCPPPLADHLNSTNH 354  
QY 401 AIQIQLMNSMDPGSTCCVPTKLTPIISILYIDAGNNVNYKYQYEDVMVYSGCR 455  
DB 355 AIVQTLVNSV-SNIPKACCVPTLSAISMLYLDVYKVLKNYQEMVYEGGCR 408  
RESULT 14  
BMP2\_HUMAN  
ID BMP2\_HUMAN STANDARD; PRT; 396 AA.  
AC P12643;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Bone morphogenetic protein 2 precursor (BMP-2) (BMP-2A).  
GN BMP2 OR BMP2A.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89072730; PubMed=3201241;  
RA Wozney J.M., Rosen V., Celeste A.J., Mittleman M.J.,  
RA Kriz R.W., Hewick R.M., Wang E.A.;  
RT "Novel regulators of bone formation: molecular clones and  
RT activities.";  
RL Science 242:1528-1534(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Shore E.M., Xu M.-Q., Calvert G., Moriatas J., Kaplan F.S.;  
RT "Human bone morphogenetic protein 2 (BMP-2) genomic DNA sequence.";  
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21638749; PubMed=11780052;  
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L.,  
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
RA Coulson A., Coville G.J., Deadman R., Dhali P.D., Dunn M.,  
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
RA Graffham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Leivasalio M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McConachie L.J., McIlroy K., McKernan A.A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,  
RA Skuce R.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,

RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:865-871(2001).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 292-396.  
 RX MEDLINE:99175323; PubMed:10074410;  
 RA Scheufler C., Sebald W., Huelsmeyer M.;  
 RT "Crystal structure of human bone morphogenetic protein-2 at 2.7 A  
 RT resolution";  
 RL J. Mol. Biol. 287:103-115(1999).  
 CC -1- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION.  
 CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.  
 CC -1- TISSUE SPECIFICITY: PARTICULARLY ABUNDANT IN LUNG, SPLEEN AND  
 CC COLON AND IN LOW BUT SIGNIFICANT LEVELS IN HEART, BRAIN, PLACENTA,  
 CC LIVER, SKELETAL MUSCLE, KIDNEY, PANCREAS, PROSTATE, OVARY AND  
 CC SMALL INTESTINE.  
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M22489; AAA51834.1; -;  
 DR EMBL; AF040249; AAF21646.1; -;  
 DR EMBL; AL035668; CAB82007.1; -;  
 DR PIR; B37278; B37278.  
 DR PDB; 3BMP; 12-MAR-00.  
 DR Genew; HGNC:1069; BMP2.  
 DR MIM; 112261; -;  
 DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb.N.  
 DR Pfam; PF000019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFb-propeptide; 1.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFb; 1.  
 DR PROSITE; PS00250; TGF-BETA\_1; 1.  
 KW Signal; Growth factor; Cytokine; Bone; Cartilage; Glycoprotein;  
 KW 3D-structure. 1 23 POTENTIAL.  
 FT SIGNAL 1 23  
 FT PROPEP 24 282  
 FT CHAIN 283 396  
 FT DISULFID 296 361  
 FT DISULFID 325 393  
 FT DISULFID 329 395  
 FT DISULFID 360 360  
 FT CARBOHYD 135 135  
 FT CARBOHYD 163 163  
 FT CARBOHYD 164 164  
 FT CARBOHYD 200 200  
 FT SEQUENCE 396 AA; 44702 MW; 20653A3987B25E60 CRC64;  
 Query Match 17.6%; Score 431.5; DB 1; Length 396;  
 Best Local Similarity 29.2%; Pred. No. 5.1e-22;  
 Matches 130; Conservative 67; Mismatches 139; Indels 109; Gaps 20;  
 QY 49 GKWRAPRDSADGREGEPQPRQDEPRAQPP-----RAQEPGGRPVVPHEYMLS 100  
 DB 23 GLVPELGRKFAAASSGRFSSQSDSELFELRLLSMFGKLRKQTPPSRDVVP-PYMLD 81  
 QY 101 IYRTSYAEKLGINAS-----PFQSKSANTITSF-VDRGLDLSHTPLR-RQKYLFDV 152  
 DB 82 LYRHS-----GQPGFADPHRLERASRANTVRSFHEESLEELPSTKTRRRFFNL 136  
 QY 153 SMLSDKEELYGLRFLRQAPSAPGPPAGPLH-VOLFCLSP-----LILDAT 201  
 DB 137 SSPTTEFFITSAELOVFEQMDALGNSSFFHRIYELIKPATANSKFPVTRLLDTRL 196  
 QY 202 LDPQAGPAGWEVDVWQGLRHQFWKQLCLELRANGELDAGEAEARARGQPPPPDLR 261

DB 197 VN-QNA--SRWESFDVTPAYMR--W-----TAQGHANHFVVEVAHLEEKQ----- 237  
 QY 262 SLGFGRRVRPQERALLVWTRSORKNLFAEMREQLGSAEAGPGACGAGSGWPPSPGAPD 321  
 DB 238 --GVSKR-----HVRISRLHQD-----EHSW-----SQ 259  
 QY 322 ARWLPSPG-----RRRRRTAFASRHGKRHKSRRLR--CSKKPLHVNFKELQWDD 370  
 DB 260 IRPLLVTFGHDGKGPLHKREKQA-----XHKQKRLKSSCKRHPLYVDFSDVGWDD 312  
 QY 371 WITAPLEYAYHCEGVCDFPLRSHLEPTNHAITQTLNMDPGSTPPSCCVPKLTISI 430  
 DB 313 WIWAPPGYHAFYCHGECPPFLADHLNTHAIVQTLVNSV-SKIPKACCCVPTLSAISM 371  
 QY 431 LYIDAGNNVYKQVEDMVVESCGR 455  
 DB 372 LYLDENEKVKLVKQYDMVVEGCGCR 396  
 RESULT 15  
 BMP4\_RABIT  
 ID BMP4\_RABIT STANDARD; PRT; 409 AA.  
 AC O46576;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Bone morphogenetic protein 4 precursor (BMP-4).  
 GN BMP4 OR BMP-4.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 ON NCBI\_TaxID=9986;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=New Zealand white; TISSUE=Ocular ciliary epithelium;  
 RA Wan X.L., Sears J., Chen S., Sears M.;  
 RT "Cloning and expression of BMP-2/-4 from rabbit ocular ciliary  
 RT epithelium";  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION (BY SIMILARITY).  
 CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: SECRETED INTO THE EXTRACELLULAR MATRIX.  
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF042497; AAB97467.1; -;  
 DR HSSP; P12643; 3BMP.  
 DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb.N.  
 DR Pfam; PF000019; TGF-beta; 1.  
 DR Pfam; PF00688; TGFb-propeptide; 1.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFb; 1.  
 DR PROSITE; PS00250; TGF-BETA\_1; 1.  
 KW Signal; Growth factor; Cytokine; Bone; Cartilage; Glycoprotein.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT PROPEP 20 293  
 FT CHAIN 294 409  
 FT DISULFID 309 374  
 FT DISULFID 338 406  
 FT DISULFID 342 408  
 FT DISULFID 373 373  
 FT CARBOHYD 144 144  
 FT CARBOHYD 209 209  
 FT CARBOHYD 351 351  
 FT CHAIN 294 409  
 FT DISULFID 309 374  
 FT DISULFID 338 406  
 FT DISULFID 342 408  
 FT DISULFID 373 373  
 FT CARBOHYD 144 144  
 FT CARBOHYD 209 209  
 FT CARBOHYD 351 351

```
FT CARBOHYD 366 366 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 409 AA; 46641 MW; 35557561447AD625 CRC64;

Query Match      17.6%; Score 430; DB 1; Length 409;
Best Local Similarity 27.7%; Pred. No. 6.7e-22;
Matches 134; Conservative 68; Mismatches 167; Indels 114; Gaps 18;

QY 7 LLSAVFLISFLMDLPGFQOASISSSCSSAELGSKG-MRSRKEGKMQRAPRDSDA---GR 62
Db 7 MLMVLLCQVL--LGASHASLIPETGKKVAIEIQHAGRRSGQSGHELLRFEATLLQM 64

QY 63 EGQEPQRFQDEPRAOQPPRAQPPGRVPRVHEYMLSIYRTYSIAEKL-----GINA 115
Db 65 FGLRRHPQPSKS-----AVIP-DYMRDLYRLQSGEEEEEQMPSGGLE 107

QY 116 SFFQSSKSANTITSF-VDRGLDDLSTPLRRQ-KYLFVSMLSKDEELVGAELRLFRQAP 173
Db 108 PERPASR-ANTVRSFHHEHLENIPGTSSENSAFRLFNLSSIPENEAISSAELRLFRE-- 164

QY 174 SAPWGP--PAGPLHVOLFCLSP-----LLLDARTLDPOGAPPAGWEFVDVWOG 220
Db 165 QVDQGPDWERGFHRINIYEVMPPPAEAVPGHLITRLDTRLVHHN---VTRWETFDV--- 218

QY 221 LRHPQWKQLCLELRAAWGELDAGEAEARARGPQQPPPPDLRLSLGFGRRVRPPQOERALLVY 280
Db 219 -----SPAVLR---WTREKQPNHGLAVEVT 240

QY 281 FTRSQKNLFAEMREQLGSAEAGCPGAGAGSWPPSPGAPDARPWLPSGR-----R 332
Db 241 HFHTTRTHQGQHVRLSRSLQSGDWA-----QFRPLLVTFGHDGRGHALT 287

QY 333 RRTAFASRHRGKRKKSRLRCSKKPLHYNFKELGWDDWIIAPLEYEAYHCEGVCDFPLR 392
Db 288 RRAKSLAKHHQPQARKKKNKCRRHLYVDFSDVGWNDWIVAPPGYQAFYCHGDCFFPLA 347

QY 393 SHLEPTNHAIQTLMNSMDPGSTPPSCCVPTKLTPIISILIYDAGNNVYKQYEDMVVESC 452
Db 348 DHFNSTNHAIVQTLVNSVN-SSIPKACCVPTELSAISMLYLDYDKVLKNYQEMVVEGC 406

QY 453 GCR 455
Db 407 GCR 409
```

Search completed: November 25, 2002, 02:56:09  
Job time : 38 secs

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 00:20:15 : Search time 3689 Seconds  
(without alignments)  
11241.935 Million cell updates/sec

Title: US-09-825-751A-19  
Perfect score: 1425  
Sequence: 1 cccctggggagacgcagcca.....tttccgcgcgttgccgcg 1425

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pi.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vl.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pi.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vl.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rtd.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1425	100.0	1425	6	AX277361	Sequence
2	1425	100.0	1425	6	AX465491	Sequence
3	1421.8	99.8	2791	6	AX351025	Sequence
C 4	991	69.5	156873	9	AC015998	Homo sapi
C 5	991	69.5	230879	9	AP003465	Homo sapi
6	987	69.3	1171	6	I62379	Sequence 25
7	935	65.6	184987	2	AC126226	Papio cyn
8	928.2	65.1	1308	4	BT013661	Bos taurus
9	928.2	65.1	1308	6	AX277426	Sequence
10	885.8	62.2	143312	2	AC126238	Felis cat
C 11	798.8	56.1	219978	2	AC126232	Bos tauru
C 12	765.8	53.7	209733	10	AC058786	Mus muscu
13	755.4	53.0	326540	2	AL732476	Mus muscu
14	749.8	52.6	167312	2	AC116927	Rattus no
15	581.8	40.8	1046	6	I62382	Sequence 31
C 16	426.8	30.0	253694	2	AC084005	Homo sapi
C 17	389.8	27.4	177269	2	AC114354	Rattus no
C 18	377.6	26.5	136505	2	AC126926	Felis cat
C 19	360.2	25.3	1880	5	AF155125	Xenopus l
C 20	343.4	24.1	138738	2	AC126234	Canis fam
C 21	343	24.1	179366	2	AC114503	Rattus no
22	320.6	22.5	375	10	MMU08338	Mus musculu
C 23	296	20.8	134623	2	AC126235	Canis fam
C 24	285.8	20.1	1354	5	DRDYNAMO	X99769 D.rerio mrn
C 25	270.6	19.0	381	5	AF089086	Gallus ga
C 26	267	18.7	177514	2	AC126919	Bos tauru
27	262.4	18.4	2169	9	AF254568S2	Cercopith
28	259.2	18.2	1345	6	I62383	Sequence 33
29	259.2	18.2	206021	9	AC012065	Homo sapi
30	255	17.9	183402	2	AC126231	Bos tauru
31	253.6	17.8	2694	9	AF254567	Macaca fa
32	253	17.8	2323	9	HSU13660	Human carti
33	253	17.8	2703	6	AX3194	Sequence 1
34	253	17.8	2703	6	A59669	Sequence 2
35	253	17.8	2703	6	A60247	Sequence 1
36	253	17.8	2703	6	A79612	Sequence 2
37	253	17.8	2703	6	AR091289	Sequence
38	253	17.8	2703	6	E12603	CDNA encodi
39	253	17.8	2703	6	E14291	DNA encodin
40	253	17.8	2703	6	E15606	CDNA encodi
41	253	17.8	2703	6	HS0915	H.sapiens G
C 42	253	17.8	151903	9	HS0915	Human DNA
43	251.6	17.7	1207	6	AR137664	Sequence 3
44	251.6	17.7	1207	6	I62365	Sequence 3
45	251.4	17.6	387	6	A79614	Sequence 4

ALIGNMENTS

RESULT 1	AX277361	Sequence 19	1425 bp	DNA	linear	PAT 29-OCT-2001
LOCUS	AX277361	Sequence 19	from Patent	WO0174897.		
DEFINITION	AX277361					
ACCESSION	AX277361					
VERSION	AX277361.1	GI:16548926				
KEYWORDS	human.					
SOURCE	human.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1					
AUTHORS	Vernet,C.A., Burgess,C.E., Fernandes,E., Taupier,R.J., Quinn,K.E., Spytek,K.A., Rastelli,L. and Herrmann,J.L.					
TITLE	Novel proteins and nucleic acids encoding same					



/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
31..1398  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD4373.1"  
/db\_xref="GI:21899854"  
/translation="MDTRVLLSAVFLISFLWDLPGFOQASISSCSAELGSKGMR  
SRKEGMRAPRDSQAGREGQPPQPODEPRAQOPRQGRVPRVPHYMLSIY  
RTYSIAELGINSFQDSKANTITSFVDRLDLSHTPLRGRVLPDVSMLSKKE  
LVGAEURLFQAPSPAGPPLVPLPCLSLLDLADLPDOPAGPACWEVDFVW  
QGLRHOPKQLCLELRANGELDAEAEARARGPOQPPDLRLSFGFRVYRPOERA  
LLVFTGRQKLNLFEMREQLDGAEGPAGAEGWPPSPGAPDARWPLPSPGRRR  
RTAFASRHGRHKKRSKPLHVNFKELGWDWIAPLEYEAYHCEGVDFPLR  
SHEPTNHALIOTLMSNDPGSTPPSCVPTKLTPISILYIDAGNNVYVQYEDMVE  
SGCR"

BASE COUNT 234 a 488 c 469 g 234 t  
ORIGIN

Query Match 100.0%; Score 1425; DB 6; Length 1425;  
Best Local Similarity 100.0%; Pred. No. 3.9e-215;  
Matches 1425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCTGGGAGAGCGACCTTGCCTCCGCGCATGATCTCCAGGCTCCTGCTCTCGGCC 60  
DB 1 CTCTGGGAGAGCGACCTTGCCTCCGCGCATGATCTCCAGGCTCCTGCTCTCGGCC 60

QY 61 GTCTTCTCATAGTTTCTGTGGGATTTGCCGGTTTCCAGCAGGCTTCCATCTCATCC 120  
DB 61 GTCTTCTCATAGTTTCTGTGGGATTTGCCGGTTTCCAGCAGGCTTCCATCTCATCC 120

QY 121 TCCTGTTCTGCCCGAGTGGGTTCCACCAAGGGCATCGAAGCGCGCAAGGAAGCAAG 180  
DB 121 TCCTGTTCTGCCCGAGTGGGTTCCACCAAGGGCATCGAAGCGCGCAAGGAAGCAAG 180

QY 181 ATGACGGGGCGCGCGACAGTACGCGGGCGGGAGGGCGAGAACACACAGCGCGG 240  
DB 181 ATGACGGGGCGCGCGACAGTACGCGGGCGGGAGGGCGAGAACACACAGCGCGG 240

QY 241 CCTCAGAGCAACCCCGGCTCAGCAGCCCGGGCGCAGAGCGCGCAGGAGGCTCCG 300  
DB 241 CCTCAGAGCAACCCCGGCTCAGCAGCCCGGGCGCAGAGCGCGCAGGAGGCTCCG 300

QY 301 CGCGTGGTCCCGCACGAGTACATGTGTCAATCTACAGGACTTACTCCATCGGTGAGAAG 360  
DB 301 CGCGTGGTCCCGCACGAGTACATGTGTCAATCTACAGGACTTACTCCATCGGTGAGAAG 360

QY 361 CTGGGATCAATGCCAGCTTTTCCAGTCTTCCAGTCCGCTCAATACGATCACCAGCTTT 420  
DB 361 CTGGGATCAATGCCAGCTTTTCCAGTCTTCCAGTCCGCTCAATACGATCACCAGCTTT 420

QY 421 GTAGACAGGGGACTAGACGATCTCTCGACACTCTCTCCGGAGACAGAAATTTTGTTT 480  
DB 421 GTAGACAGGGGACTAGACGATCTCTCGACACTCTCTCCGGAGACAGAAATTTTGTTT 480

QY 481 GATGTCTCATGCTCTCAGACAAAGAGAGCTGTGGGGCGGAGGCTCGGCTCTTTTCGC 540  
DB 481 GATGTCTCATGCTCTCAGACAAAGAGAGCTGTGGGGCGGAGGCTCGGCTCTTTTCGC 540

QY 541 CAGGCGCCCTCAGCGCTTGGGGCCACACAGCGGGCGGCTCCACGTGAGCTCTTCCT 600  
DB 541 CAGGCGCCCTCAGCGCTTGGGGCCACACAGCGGGCGGCTCCACGTGAGCTCTTCCT 600

QY 601 TGCTTTTCGCCCTTACTGCTGGACGCGCGGACCTTGGACCCGAGGGCGCGCGGCC 660  
DB 601 TGCTTTTCGCCCTTACTGCTGGACGCGCGGACCTTGGACCCGAGGGCGCGCGGCC 660

QY 661 GGCTGGGAAGTCTTCGACGTGTGGCAGGCGCTCGCGCACAGCCCTGGAAGCAGCTGTGC 720  
DB 661 GGCTGGGAAGTCTTCGACGTGTGGCAGGCGCTCGCGCACAGCCCTGGAAGCAGCTGTGC 720

QY 721 TTGGAGCTCGGGCGGCATGGGCGAGCTGGAGCGCGGGAGCGCGCGCGCGCG 780  
DB 721 TTGGAGCTCGGGCGGCATGGGCGAGCTGGAGCGCGGGAGCGCGCGCGCGCGCG 780

Db 721 TTGGAGCTCGGGCGGCATGGGCGAGCTGGACCGCGGAGCGCGCGCGCGG 780  
QY 781 GGACCCAGCAACCGCGCGCGCGCGAGCTGCGGAGTCTGGGCTTTCGGCGGAGGTCGG 840  
Db 781 GGACCCAGCAACCGCGCGCGCGCGAGCTGCGGAGTCTGGGCTTTCGGCGGAGGTCGG 840  
QY 841 CTCTCCAGAGAGCGGCGCTGCTGTGGTATTCACAGATCCCGAGCGCAAGAACCTGTT 900  
Db 841 CTCTCCAGAGAGCGGCGCTGCTGTGGTATTCACAGATCCCGAGCGCAAGAACCTGTT 900  
QY 901 GCAGATGCGCGAGCAGCTGGGCTCGGCGAGGCTCGGCGCGCGCGCGCGCGCGAG 960  
Db 901 GCAGATGCGCGAGCAGCTGGGCTCGGCGAGGCTCGGCGCGCGCGCGCGCGCGAG 960  
QY 961 GGCTGCTGGCGCGCGCTGGGCGCGCGGATGCCAGGCTTGGCTGCCCTCGCGCGCG 1020  
Db 961 GGCTGCTGGCGCGCGCTGGGCGCGCGGATGCCAGGCTTGGCTGCCCTCGCGCGCG 1020  
QY 1021 CGCGCGCGCGCGCGCGCTTTCGCGAGTGCCTATGCGAAGCGCGCACGCAAGATCC 1080  
Db 1021 CGCGCGCGCGCGCGCGCTTTCGCGAGTGCCTATGCGAAGCGCGCACGCAAGATCC 1080  
QY 1081 AGCTAGCTGCGAGCAAGAGCCCTGCGAGTGAATTCAGAGCTGGCTGGGAGGAC 1140  
Db 1081 AGCTAGCTGCGAGCAAGAGCCCTGCGAGTGAATTCAGAGCTGGCTGGGAGGAC 1140  
QY 1141 TGGATTATCGCGCGCTGGAGTACGAGGCTTACTGCGAGGCTGTATGCGACTTCCCG 1200  
Db 1141 TGGATTATCGCGCGCTGGAGTACGAGGCTTACTGCGAGGCTGTATGCGACTTCCCG 1200  
QY 1201 CTGCGTCTGCGACCTGGAGCCCAACACCGCCATCATCCAGCGCTGATGAATCCATG 1260  
Db 1201 CTGCGTCTGCGACCTGGAGCCCAACACCGCCATCATCCAGCGCTGATGAATCCATG 1260  
QY 1261 GACCCGCGCTCCACCCCGCGCGCTGCTGGTGGTCCCGACCAATTTGACTCCCATCAGATT 1320  
Db 1261 GACCCGCGCTCCACCCCGCGCGCTGCTGGTGGTCCCGACCAATTTGACTCCCATCAGATT 1320  
QY 1321 CTATACATCGACGCGGCAATATGTTGCTTACAAGCAGTACGAGGACATGGTGGAG 1380  
Db 1321 CTATACATCGACGCGGCAATATGTTGCTTACAAGCAGTACGAGGACATGGTGGAG 1380  
QY 1381 TCGTGGGTGCGAGTAGCGGTGCTTCCCGCGCTTGGCGCG 1425  
Db 1381 TCGTGGGTGCGAGTAGCGGTGCTTCCCGCGCTTGGCGCG 1425

RESULT 3  
AX351025 2791 bp DNA linear PAT 06-FEB-2002  
DEFINITION Sequence 15 from Patent WO0194587.  
ACCESSION AX351025  
VERSION AX351025.1 GI:18616381  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Lal,P., Yue,H., He,A., Nguyen,D.B., Walla,N., Gandhi,A.R., Azimza,Y., Bandman,O., Tang,Y.T., Lu,Y., Baughn,M.R., Duggan,B.M., Lee,S., Hafalia,A. and Policky,J.L.  
TITLE Extracellular messengers  
JOURNAL Patent: WO 0194587-A 15 13-DEC-2001;  
FEATURES Incyte Genomics, Inc. (US)  
Location/Qualifiers  
1..2791  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 7481107CB1"  
BASE COUNT 454 a 945 c 942 g 450 t  
ORIGIN

Query Match 99.8%; Score 1421.8; DB 6; Length 2791; Best Local Similarity 99.9%; Pred. No. 1.le-214; Matches 1423; Conservative 0; Mismatches 2; Indels 0; Gaps 0;									
QY	1	CTCTGGGAGAGCGACGACCTTGGCCCGCCATGGATACCTCCAGGGTCTCTCTCGGGCC	60						
Db	999	CTCCTGGGAGAGCGACGACCTTGGCCCGCCATGGATACCTCCAGGGTCTCTCTCGGGCC	1058						
QY	61	GTCTTCTCATCAGTTTCTGTGGATTTGCCGGTTTCCAGCAGGCTTCCATCTCATCC	120						
Db	1059	GTCTTCTCATCAGTTTCTGTGGATTTGCCGGTTTCCAGCAGGCTTCCATCTCATCC	1118						
QY	121	TCCTGTTCGTCGCCGAGCTGGGTTCACCAAGGGCATCGAAGCCGCAAGGAAGCAAG	180						
Db	1119	TCCTGTTCGTCGCCGAGCTGGGTTCACCAAGGGCATCGAAGCCGCAAGGAAGCAAG	1178						
QY	181	ATGCAGGGGCGCGCGCGACAGTACGCGGGCGGGAGGGCCAGGAACACAGCGCGGG	240						
Db	1179	ATGCAGGGGCGCGCGCGACAGTACGCGGGCGGGAGGGCCAGGAACACAGCGCGGG	1238						
QY	241	CCTCAGGACGAACCCCGGGCTCAGCAGCCCGGGCGCAGGAGCCGCCAGGAGGTCCG	300						
Db	1239	CCTCAGGACGAACCCCGGGCTCAGCAGCCCGGGCGCAGGAGCCGCCAGGAGGTCCG	1298						
QY	301	CGCGTGGTGCCCGACAGTACATGTGTCAATCTACAGGACTTACTTCCATCGTGAGAAG	360						
Db	1299	CGCGTGGTGCCCGACAGTACATGTGTCAATCTACAGGACTTACTTCCATCGTGAGAAG	1358						
QY	361	CTGGGATCAATGCCAGCTTTTTCAGTCTTCCAGTTCGGCTTAATACGATCACCAGCTTT	420						
Db	1359	CTGGGATCAATGCCAGCTTTTTCAGTCTTCCAGTTCGGCTTAATACGATCACCAGCTTT	1418						
QY	421	GTAGCAGGGGACTAGACGATCTCTGCACACTCCTCTCCGGAGACAGAAGTATTGTGTT	480						
Db	1419	GTAGCAGGGGACTAGACGATCTCTGCACACTCCTCTCCGGAGACAGAAGTATTGTGTT	1478						
QY	481	GATGTGTCCATGCTCTCAGACAAAGAAGAGCTGTGTGGCGCGGAGTGGCGTCTTTTCGC	540						
Db	1479	GATGTGTCCATGCTCTCAGACAAAGAAGAGCTGTGTGGCGCGGAGTGGCGTCTTTTCGC	1538						
QY	541	CAGGCGCCCTCAGCGCCCTGGGGGACACAGCGGGCGGCTCCAGTGCAGCTCTTCCCT	600						
Db	1539	CAGGCGCCCTCAGCGCCCTGGGGGACACAGCGGGCGGCTCCAGTGCAGCTCTTCCCT	1598						
QY	601	TGCCTTTTCGCCCTACTGCTGGACGCGCGGACCCCTGGACCCCGAGGGCGCGCGGCC	660						
Db	1599	TGCCTTTTCGCCCTACTGCTGGACGCGCGGACCCCTGGACCCCGAGGGCGCGCGGCC	1658						
QY	661	GGCTGGGAAGTCTTCGACGTGTGGCAGGGCCCTGCGCCACACGACCCCTGGAAGCAGCTGTGC	720						
Db	1659	GGCTGGGAAGTCTTCGACGTGTGGCAGGGCCCTGCGCCACACGACCCCTGGAAGCAGCTGTGC	1718						
QY	721	TTGGAGCTCGGGCCGATGGGGCAGCTGAGCGCGGGGAGCGCGAGCGCGCGGGG	780						
Db	1719	TTGGAGCTCGGGCCGATGGGGCAGCTGAGCGCGGGGAGCGCGAGCGCGCGGGG	1778						
QY	781	GGACCCCGACACCGCCGCCCGGACCTGCGGAGTCTGGGCTTGGCGGAGGCTGGG	840						
Db	1779	GGACCCCGACACCGCCGCCCGGACCTGCGGAGTCTGGGCTTGGCGGAGGCTGGG	1838						
QY	841	CCTCCCCAGGAGCGGGCCCTGCTGTGTGTATTACACAGATCCACGCAAGAACCTGTTC	900						
Db	1839	CCTCCCCAGGAGCGGGCCCTGCTGTGTGTATTACACAGATCCACGCAAGAACCTGTTC	1898						
QY	901	GCAGAGATCGGGAGCAGCTGGGCTTGGCCGAGGCTGGGGCCCCGGGCGCGCGCGAG	960						
Db	1899	GCAGAGATCGGGAGCAGCTGGGCTTGGCCGAGGCTGGGGCCCCGGGCGCGCGCGAG	1958						
QY	961	GGGTCTGTGCCCGCCGCTGGCGGCCCGCGGATGCCAGGCTTGGCTGCGCCCTCGCGCG	1020						
Db	1959	GGGTCTGTGCCCGCCGCTGGCGGCCCGCGGATGCCAGGCTTGGCTGCGCCCTCGCGCG	2018						
QY	1021	CGCCGGCGGGCGCACGGGCCCTTCGCCAGTGCCTATGGCAGCGGCACGGAAGTCC	1080						
Db	2079	CGCGGGCGGGCGCACGGGCCCTTCGCCAGTGCCTATGGCAGCGGCACGGAAGTCC	2078						
QY	1081	AGGCTACGCTGCAGCAAGAAGCCCTGACGTGAACCTTCAAGAGAGTGGGCTGGGACGAC	1140						
Db	2079	AGGCTACGCTGCAGCAAGAAGCCCTGACGTGAACCTTCAAGAGAGTGGGCTGGGACGAC	2138						
QY	1141	TGGATTATCGCGCCCTGGAGTACGAGGCTATCACTGCGAGGGTGTATGCCACTTCCCG	1200						
Db	2139	TGGATTATCGCGCCCTGGAGTACGAGGCTATCACTGCGAGGGTGTATGCCACTTCCCG	2198						
QY	1201	CTGCGCTCGACCTGAGAGCCCAACACGCGCATCATCAGACGCTGATGAATCCATG	1260						
Db	2199	CTGCGCTCGACCTGAGAGCCCAACACGCGCATCATCAGACGCTGATGAATCCATG	2258						
QY	1261	GACCCGGGCTCACCCCGCCAGCTGCTGCGTGCGCCACCAATTTGACTCCCATCAGCAT	1320						
Db	2259	GACCCGGGCTCACCCCGCCAGCTGCTGCGTGCGCCACCAATTTGACTCCCATCAGCAT	2318						
QY	1321	CTATACATCGACGCGGGCAATATGTGTCTACAAGCATGAGGACATGGTGGTGAG	1380						
Db	2319	CTATACATCGACGCGGGCAATATGTGTCTACAAGCATGAGGACATGGTGGTGAG	2378						
QY	1381	TCGTGCGGCTGCAGGTAGCGGTGCTTTCCCGCGGCTTTGGCCCG	1425						
Db	2379	TCGTGCGGCTGCAGGTAGCGGTGCTTTCCCGCGGCTTTGGCCCG	2423						
RESULT 4									
AC015998/c									
LOCUS Homo sapiens, clone RP11-15J14, 156873 bp DNA linear PRI 01-MAY-2001									
DEFINITION AC015998									
ACCESSION AC015998									
VERSION AC015998.8 GI:13399387									
KEYWORDS HTG.									
SOURCE Homo sapiens.									
ORGANISM Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE 1 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,									
TITLE Homo sapiens, clone RP11-15J14									
JOURNAL Unpublished									
REFERENCE 2 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,									
TITLE Homo sapiens, clone RP11-15J14									
JOURNAL Unpublished									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									
JOURNAL Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome									
REFERENCE 3 (bases 1 to 156873)									
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,									
TITLE Direct Submission									



Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Sounez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (01-MAY-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 21, 2001 this sequence version replaced gi:11560231.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE  
JOURNAL  
COMMENT

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L3477  
Center clone name: I5\_J\_14  
-----

FEATURES  
source

Location/Qualifiers  
1. .156873  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="RP11-15J14"  
/clone\_lib="RPC1-11 Human Male BAC"  
1. .87

repeat_region	/rpt_family="MLT1F1"	/rpt_family="L2"
repeat_region	108. .158	complement(8208. .8688)
repeat_region	200. .274	/rpt_family="L2"
repeat_region	/rpt_family="L2"	8850. .8876
repeat_region	1372. .1490	/rpt_family="(CATA)n"
repeat_region	1493. .1539	/rpt_family="L1MB3"
repeat_region	/rpt_family="(TTA)n"	9097. .9097
repeat_region	1599. .1763	9098. .9407
repeat_region	/rpt_family="FRAM"	/rpt_family="AluJb"
repeat_region	complement(2665. .2732)	9408. .9498
repeat_region	2747. .2873	/rpt_family="L1MB3"
repeat_region	/rpt_family="L2"	9408. .9498
repeat_region	3041. .3081	/rpt_family="L1MA4A"
repeat_region	/rpt_family="L2"	9588. .10083
repeat_region	3446. .3472	/rpt_family="L1MA4A"
repeat_region	3476. .3501	10084. .10129
repeat_region	/rpt_family="AT_rich"	/rpt_family="(CAAAA)n"
repeat_region	/rpt_family="(TTTAA)n"	10130. .10892
repeat_region	3807. .3911	/rpt_family="L1MA4A"
repeat_region	/rpt_family="MER94"	10893. .11189
repeat_region	complement(4180. .4328)	/rpt_family="AluSx"
repeat_region	complement(5606. .5809)	11190. .11413
repeat_region	/rpt_family="L1P5"	/rpt_family="L1MA4A"
repeat_region	5866. .6149	11592. .11845
repeat_region	/rpt_family="AluSq"	/rpt_family="L1MB3"
repeat_region	6332. .6461	11886. .11916
repeat_region	6492. .6595	/rpt_family="AT_rich"
repeat_region	/rpt_family="FLAM_C"	11920. .12153
repeat_region	6811. .7150	/rpt_family="L1MA9"
repeat_region	/rpt_family="L1MB3"	12154. .12175
repeat_region	7425. .7453	/rpt_family="(TAAA)n"
repeat_region	/rpt_family="AT_rich"	/rpt_family="L1MA9"
repeat_region	7699. .7917	12176. .12328
repeat_region	complement(7920. .8214)	/rpt_family="L1MA9"
repeat_region		12329. .12362
repeat_region		/rpt_family="(TAAA)n"
repeat_region		12363. .12670
repeat_region		/rpt_family="L1MA9"
repeat_region		complement(12671. .13025)
repeat_region		/rpt_family="THE1A"
repeat_region		13026. .13669
repeat_region		/rpt_family="L1MA9"
repeat_region		13711. .14022
repeat_region		/rpt_family="AluSx"
repeat_region		14082. .14206
repeat_region		/rpt_family="L1MB3"
repeat_region		15453. .15532
repeat_region		/rpt_family="L1MB3"
repeat_region		16551. .16618
repeat_region		/rpt_family="L2"
repeat_region		16730. .16756
repeat_region		/rpt_family="AT_rich"
repeat_region		17583. .17891
repeat_region		/rpt_family="AluSx"
repeat_region		17814. .17942
repeat_region		/note="Less than 30 qual SNGL region"
repeat_region		17901. .18037
repeat_region		/rpt_family="GA-rich"
repeat_region		complement(18164. .18466)
repeat_region		/rpt_family="AluY"
repeat_region		complement(19025. .19239)
repeat_region		/rpt_family="MER20"
repeat_region		19526. .19564
repeat_region		/rpt_family="(CAAAA)n"
repeat_region		19917. .19938
repeat_region		/rpt_family="AT_rich"
repeat_region		20711. .20821
repeat_region		/rpt_family="MIR3"
repeat_region		21514. .22019
repeat_region		/rpt_family="L2"
repeat_region		22107. .22209
repeat_region		/rpt_family="L2"
repeat_region		complement(22210. .22330)
repeat_region		/rpt_family="MLT1A1"
repeat_region		22334. .22989
repeat_region		/rpt_family="L1"

[illegible]

/evidence-not\_experimental  
/rpt\_family="L2"  
9881. .9915  
/evidence-not\_experimental  
/rpt\_family="(A)n"  
9987. .10233  
/evidence-not\_experimental  
/rpt\_family="MIR"  
/complement(10279. .10451)  
/evidence-not\_experimental  
/rpt\_family="MER20"  
10823. .10869  
/evidence-not\_experimental  
/rpt\_family="L2"  
14352. .14387  
/evidence-not\_experimental  
/rpt\_family="(TCAA)n"  
/complement(14742. .15115)  
/evidence-not\_experimental  
/rpt\_family="MLTlk"  
15956. .15976  
/evidence-not\_experimental  
/rpt\_family="(TC)n"  
16580. .16624  
/evidence-not\_experimental  
/rpt\_family="L2"  
17179. .17309  
/evidence-not\_experimental  
/rpt\_family="MER94"  
18483. .18503  
/evidence-not\_experimental  
/rpt\_family="GC\_rich"  
18730. .18886  
/evidence-not\_experimental  
/rpt\_family="MIR"  
19782. .20109  
/evidence-not\_experimental  
/rpt\_family="MLTlj"  
/complement(20367. .20418)  
/evidence-not\_experimental  
/rpt\_family="MER47"  
20423. .20482  
/evidence-not\_experimental  
/rpt\_family="MLTlj"  
20489. .20523  
/evidence-not\_experimental  
/rpt\_family="(T)n"  
20937. .21009  
/evidence-not\_experimental  
/rpt\_family="L2"  
21112. .21141  
/evidence-not\_experimental  
/rpt\_family="(TAA)n"  
21152. .21178  
/evidence-not\_experimental  
/rpt\_family="(TTTTC)n"  
/complement(21181. .21298)  
/evidence-not\_experimental  
/rpt\_family="AluJo/FRAM"  
21328. .21634  
/evidence-not\_experimental  
/rpt\_family="MER2"  
23969. .24028  
/evidence-not\_experimental  
/rpt\_family="MIR"  
24846. .25168  
/evidence-not\_experimental  
/rpt\_family="Alusx"  
/complement(25610. .25702)  
/evidence-not\_experimental  
/rpt\_family="L2"  
26885. .27112  
/evidence-not\_experimental

repeat\_region  
/rpt\_family="MIR"  
29490. .29536  
/evidence-not\_experimental  
/rpt\_family="(GA)n"  
30171. .30193  
/evidence-not\_experimental  
/rpt\_family="GC\_rich"  
30215. .30269  
/evidence-not\_experimental  
/rpt\_family="GC\_rich"  
31101. .31412  
/evidence-not\_experimental  
/rpt\_family="L2"  
31927. .31953  
/evidence-not\_experimental  
/rpt\_family="(CA)n"  
31953. .31979  
/evidence-not\_experimental  
/rpt\_family="(GA)n"  
/complement(32802. .32978)  
/evidence-not\_experimental  
/rpt\_family="MIR"  
/complement(33702. .33866)  
/evidence-not\_experimental  
/rpt\_family="MER5A"  
34839. .34884  
/evidence-not\_experimental  
/rpt\_family="LIMA8"  
34890. .34916  
/evidence-not\_experimental  
/rpt\_family="AT\_rich"  
/complement(34954. .35023)  
/evidence-not\_experimental  
/rpt\_family="MIR"  
/complement(35705. .35923)  
/evidence-not\_experimental  
/rpt\_family="MIR"  
/complement(36531. .36718)  
/evidence-not\_experimental  
/rpt\_family="MIR"  
37008. .37226  
/evidence-not\_experimental  
/rpt\_family="MIR"  
39548. .39568  
/evidence-not\_experimental  
/rpt\_family="GC\_rich"  
41600. .41706  
/evidence-not\_experimental  
/rpt\_family="(TC)n"  
42168. .42400  
/evidence-not\_experimental  
/rpt\_family="MIR"  
44440. .44511  
/evidence-not\_experimental  
/rpt\_family="CT\_rich"  
45284. .45340  
/evidence-not\_experimental  
/rpt\_family="GA\_rich"  
46113. .46158  
/evidence-not\_experimental  
/rpt\_family="G-rich"  
46237. .46291  
/evidence-not\_experimental  
/rpt\_family="GC\_rich"  
46918. .47077  
/evidence-not\_experimental  
/rpt\_family="CT\_rich"  
/complement(47343. .47520)  
repeat\_region

Query Match 69.5%; Score 991; DB 9; Length 230879;  
Best Local Similarity 100.0%; Pred. No. 4e-147;  
Matches 991; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 435 AGACGATCTCTCGCACACTCTCTCCGGAGACAGAGTATTTGTTGATGTGTCCATGCT 494  
Db 30771 AGACGATCTCTCGCACACTCTCTCCGGAGACAGAGTATTTGTTGATGTGTCCATGCT 30712  
QY 495 CTCAGACAAAGAGCTGGTGGGCGCGAGCTGGGCTCTTTTCGCCAGCGCCCTCAGC 554  
Db 30711 CTCAGACAAAGAGCTGGTGGGCGCGAGCTGGGCTCTTTTCGCCAGCGCCCTCAGC 30652  
QY 555 GCCCTGGGGCCACACAGCGCGCGCTCCACCTGCAGCTCTCCCTTGCTTTCGCCCT 614  
Db 30651 GCCCTGGGGCCACACAGCGCGCGCTCCACCTGCAGCTCTCCCTTGCTTTCGCCCT 30592  
QY 615 ACTGTGGACGGCGGACCTTGACCCGAGGGGGCGCGCGCGGCTGGGAAAGTCTT 674  
Db 30591 ACTGTGGACGGCGGACCTTGACCCGAGGGGGCGCGCGCGGCTGGGAAAGTCTT 30532  
QY 675 CGACGTGGGAGGGCTGGCCACACAGCCCTGGAGCAGCTGTGTTGGAGCTGGGGC 734  
Db 30531 CGACGTGGGAGGGCTGGCCACACAGCCCTGGAGCAGCTGTGTTGGAGCTGGGGC 30472  
QY 735 CGCATGGGCGAGCTGGACCGCGGGAGCGCGAGCGCGCGGGGACCCAGCAACC 794  
Db 30471 CGCATGGGCGAGCTGGACCGCGGGAGCGCGAGCGCGCGGGGACCCAGCAACC 30412  
QY 795 GCCGCCCGGACCTGCGGAGTCTGGGCTTCGGCGGGAGGGTCCGCCCTCCCCAGGAGCG 854  
Db 30411 GCCGCCCGGACCTGCGGAGTCTGGGCTTCGGCGGGAGGGTCCGCCCTCCCCAGGAGCG 30352  
QY 855 GGCCTGCTGTGTTATTCACCATGATCCAGCGCAAGAACCTGTTGCGAGAGATCGCGA 914  
Db 30351 GGCCTGCTGTGTTATTCACCATGATCCAGCGCAAGAACCTGTTGCGAGAGATCGCGA 30292  
QY 915 GCAGCTGGGCTCGGCGAGGCTCGGGCCCGCGCGCGCGAGGGTCTGTGGCCGCG 974  
Db 30291 GCAGCTGGGCTCGGCGAGGCTCGGGCCCGCGCGCGCGAGGGTCTGTGGCCGCG 30232  
QY 975 GCCGTGGGCGCCCGGATGCGAGGCTTGGCTGCCCTCGCCCGCGCGCGCGCGCGCG 1034  
Db 30231 GCCGTGGGCGCCCGGATGCGAGGCTTGGCTGCCCTCGCCCGCGCGCGCGCGCGCG 30172  
QY 1035 CACGGCTTCGCGAGTGCATGGCAAGCGGCACGGCAAGAGTCCAGGCTACGCTGCAG 1094  
Db 30171 CACGGCTTCGCGAGTGCATGGCAAGCGGCACGGCAAGAGTCCAGGCTACGCTGCAG 30112  
QY 1095 CAAGAAGCCCTGCAGCTGAACCTCAAGAGCTGGGCTGGGAGCTGATTCGCGCG 1154  
Db 30111 CAAGAAGCCCTGCAGCTGAACCTCAAGAGCTGGGCTGGGAGCTGATTCGCGCG 30052  
QY 1155 CCGTGGTACGAGGCTATCTACTGCGAGGGTGTATGCGACTTCCGCTCGCTCGCACT 1214  
Db 30051 CCGTGGTACGAGGCTATCTACTGCGAGGGTGTATGCGACTTCCGCTCGCTCGCACT 29992  
QY 1215 GGAGCCCAACACCGCCATCATCAGACGCTGATGAATCCATGGACCCCGGCTCCAC 1274  
Db 29991 GGAGCCCAACACCGCCATCATCAGACGCTGATGAATCCATGGACCCCGGCTCCAC 29932  
QY 1275 CCGCGCCAGCTGCTGGTGGCCACCAAAATTGACTCCCATCAGCTTCTATACATCGACCG 1334  
Db 29931 CCGCGCCAGCTGCTGGTGGCCACCAAAATTGACTCCCATCAGCTTCTATACATCGACCG 29872  
QY 1335 GGGCAATATGTGTTCTACAAGCAGTACGAGACATGTTGGTGGAGTCTGGGCTGCGAG 1394  
Db 29871 GGGCAATATGTGTTCTACAAGCAGTACGAGACATGTTGGTGGAGTCTGGGCTGCGAG 29812  
QY 1395 GTAGCGGTGCTTTCCGCGCGCTTTGGCCCG 1425  
Db 29811 GTAGCGGTGCTTTCCGCGCGCTTTGGCCCG 29781

RESULT 6  
162379  
LOCUS  
DEFINITION Sequence 25 from patent US 5658882.  
1171 bp DNA linear PAT 07-OCT-1997

ACCESSION I62379  
VERSION I62379.1 GI:2480327  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1171)  
AUTHORS Celeste,A.J., Wozney,J.M., Rosen,V.A., Wolfman,N.M., Thomsen,G.H.  
and Melton,D.A.  
TITLE Methods of inducing formation of tendon and/or ligament tissue  
comprising administering BMP-12, BMP-13, and/or MP-52  
JOURNAL Patent: US 5658882-A 25 19-AUG-1997;  
FEATURES Location/Qualifiers  
1. 1171  
Source /organism="unknown"  
BASE COUNT 195 a 384 c 413 g 179 t  
ORIGIN  
Query Match 69.3%; Score 987; DB 6; Length 1171;  
Best Local Similarity 100.0%; Pred. No. 4.le-146;  
Matches 987; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 439 GATCTCTGCACACTCTCTCTCCGGAGACAGAAATTTGTTTGTATGTGTCCATGCTCA 498  
Db 8 GATCTCTGCACACTCTCTCTCCGGAGACAGAAATTTGTTTGTATGTGTCCATGCTCA 67  
QY 499 GACAAGAGAGCTGGTGGGCGGGAGCTGCGGCTCTTTTCGCCAGCGCCCTCAGCGCC 558  
Db 68 GACAAGAGAGCTGGTGGGCGGGAGCTGCGGCTCTTTTCGCCAGCGCCCTCAGCGCC 127  
QY 559 TGGGGGCCACACAGCGCGGCGCTCCACGTGCAGCTCTTCCCTTGGCTTTCGCCCTACTG 618  
Db 128 TGGGGGCCACACAGCGCGGCGCTCCACGTGCAGCTCTTCCCTTGGCTTTCGCCCTACTG 187  
QY 619 CTGGAGCGCGGACCTTGACCCCGAGGGGGCGCGCGCGCGCTGGGAAATGCTTTCGAC 678  
Db 188 CTGGAGCGCGGACCTTGACCCCGAGGGGGCGCGCGCGCGCTGGGAAATGCTTTCGAC 247  
QY 679 GTGTGGCAGGGCTTCGCCACACAGCCCTGGAAGCAGCTGTGTTGGAGCTGGGGCGCA 738  
Db 248 GTGTGGCAGGGCTTCGCCACACAGCCCTGGAAGCAGCTGTGTTGGAGCTGGGGCGCA 307  
QY 739 TGGGGCGAGCTGGAGCCCGGGAGCGCGCGCGCGCGGGAGCCCGCAGCAACCGCGG 798  
Db 308 TGGGGCGAGCTGGAGCCCGGGAGCGCGCGCGCGGGAGCCCGCAGCAACCGCGG 367  
QY 799 CCCCAGGACCTCGGAGCTTGGGCTTCGCCCGAGGGTGGGCTCCCGAGAGCGGGCC 858  
Db 368 CCCCAGGACCTCGGAGCTTGGGCTTCGCCCGAGGGTGGGCTCCCGAGAGCGGGCC 427  
QY 859 CTGCTGGTGGTATTCACAGATCCAGCGCAAGAACCTGTTTCGAGAGATGCGCAGCAG 918  
Db 428 CTGCTGGTGGTATTCACAGATCCAGCGCAAGAACCTGTTTCGAGAGATGCGCAGCAG 487  
QY 919 CTGGGCTCGGCGAGGCTTCGCCCGCGGGCGCGCGCGGGGTCGTGGCGCGCGCG 978  
Db 488 CTGGGCTCGGCGAGGCTTCGCCCGCGGGCGCGCGCGGGGTCGTGGCGCGCGCG 547  
QY 979 TCGGGCGCGCGGATGCGAGGCTTGGCTTCGCCCGCGGGCGCGCGGGCGGGCGCAG 1038  
Db 548 TCGGGCGCGCGGATGCGAGGCTTGGCTTCGCCCGCGGGCGCGCGGGCGGGCGCAG 607  
QY 1039 GCCTTCGCCAGTCGCAATGGCAAGCGGCACGCAAGAGTCCAGGCTACGCTGCAGCAAG 1098  
Db 608 GCCTTCGCCAGTTCGCAATGGCAAGCGGCACGCAAGAGTCCAGGCTACGCTGCAGCAAG 567  
QY 1099 AAGCCCTTCAGCTGAACCTTCAAGGAGCTGGGCTGGGAGCTGGATTCGCGCCCTG 1158  
Db 668 AAGCCCTTCAGCTGAACCTTCAAGGAGCTGGGCTGGGAGCTGGATTCGCGCCCTG 727  
QY 1159 GAGTACGAGGCTATCACTGCGAGGGTGTATGCGACTTCCCGCTCGGCTCGCACCTGGAG 1218  
Db 728 GAGTACGAGGCTATCACTGCGAGGGTGTATGCGACTTCCCGCTCGGCTCGCACCTGGAG 787

QY 1219 CCACCAACACCCATCATCCAGACGCTGATGAATCTCCATGACCCCGGCTCCACCCCG 1278  
|||||  
Db 788 CCCACCAACACCCATCATCCAGACGCTGATGAATCTCCATGACCCCGGCTCCACCCCG 847  
QY 1279 CCAGCTGCTGCTGCCACCAATATGACTCCATCAGATTCATATACATGACGGGGC 1338  
|||||  
Db 848 CCAGCTGCTGCTGCCACCAATATGACTCCATCAGATTCATATACATGACGGGGC 907  
QY 1339 AATAATGTGTCTACAGAGTACGAGGACATGCTGTGTGAGTCGTGGCTCAGGTAG 1398  
|||||  
Db 908 AATAATGTGTCTACAGAGTACGAGGACATGCTGTGTGAGTCGTGGCTCAGGTAG 967  
QY 1399 CGTGCTTTCCCGCGCCCTTGCCCGC 1425  
|||||  
Db 968 CGTGCTTTCCCGCGCCCTTGCCCGC 994

## RESULT 7

## AC126226

## LOCUS

AC126226 184987 bp DNA linear HTG 04-JUL-2002  
Papio cynocephalus anubis clone RP41-394G6, WORKING DRAFT SEQUENCE,  
3 unordered pieces.

## ACCESSION

## AC126226

## VERSION

## AC126226.1

## KEYWORDS

## HTG; HTGS\_PHASE1; HTGS\_DRAFT.

## SOURCE

## olive baboon.

## ORGANISM

## Papio cynocephalus anubis

## Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

## Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

## Cercopithecinae; Papio.

## 1 (bases 1 to 184987)

## Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,

## Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C.,

## Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Harshbarger, J.,

## Haghighi, P., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Laric, P.,

## Lee-Lin, S.-Q., Legaspi, R., Maduro, O.L., Maduro, V.B.,

## Margulies, E.H., Masello, C., Maskeri, B., Mastrian, S.D.,

## McCluskey, J.C., McDowell, J., Paquirigan, C., Pearson, R.,

## Portnoy, M.E., Prasad, A., Schueler, M.G., Stantripop, S., Thomas, J.W.,

## Thomas, P.J., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A.,

## Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

## NISC Comparative Sequencing Initiative

## Unpublished

## 2 (bases 1 to 184987)

## Green, E.D.

## Direct Submission

## Submitted (04-JUL-2002) NIH Intramural Sequencing Center, 8717

## Groveomont Circle, Gaithersburg, MD 20877, USA

## ----- Genome Center

## Center: NIH Intramural Sequencing Center

## Center code: NISC

Web site: <http://www.nisc.nih.gov>Contact: [nisc\\_zoo@nhgri.nih.gov](mailto:nisc_zoo@nhgri.nih.gov)

## ----- Project Information

## Center project name: der

## Center clone name: 394G06

## ----- Summary Statistics

## Sequencing vector: plasmid; n/a; 100% of reads

## Chemistry: Dye-terminator Big Dye; 100% of reads

## Assembly program: Phrap; version 0.990319

## Consensus quality: 183498 bases at least Q40

## Consensus quality: 183965 bases at least Q30

## Consensus quality: 184299 bases at least Q20

## Insert size: 177000; agarose-fp

## Insert size: 18478; sum-of-ctnigs

## Quality coverage: 9.40x in Q20 bases; agarose-fp

## Quality coverage: 9.00x in Q20 bases; sum-of-ctnigs

## -----

## \* NOTE: This is a 'working draft' sequence. It currently

## \* consists of 3 contigs. The true order of the pieces

## \* is not known and their order in this sequence record is

## \* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 40078: contig of 40078 bp in length  
\* 40079: gap of unknown length  
\* 40179: contig of 71410 bp in length  
\* 111588: gap of unknown length  
\* 111589: contig of 73299 bp in length.  
\* 111689: contig of 73299 bp in length.

## FEATURES

## source

1..184987

/organism="Papio cynocephalus anubis"

/db\_xref="taxon:9555"

/clone="RP41-394G6"

/clone\_lib="RP41"

1..40078

/note="assembly\_fragment"

clone\_end:SP6

vector\_side:right

40179..111588

/note="assembly\_fragment"

111689..184987

/note="assembly\_fragment"

clone\_end:r7

vector\_side:right

BASE COUNT 52085 a 38650 c 39949 g 54102 t 201 others

## ORIGIN

## Query Match

Best Local Similarity 65.5%; Score 935; DB 2; Length 184987;

Matches 956; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

## QY 435

AGACGATCTCTCCACACTCTCTCCGAGACAGAGTATTCTTCATGTCCTCATGCT 494

|||||

Db 178510

AGACGATCTCTCCACACTCTCTCCGAGACAGAGTATTCTTCATGTCCTCATGCT 178559

|||||

QY 495

CTCAGACAAGAAGAGCTGGTGGCGCGGAGCTCTTCGCCAGCGCCCTCAGC 554

|||||

Db 178570

CTCAGACAAGAAGAGCTGGTGGCGCGGAGCTCTTCGCCAGCGCCCTCAGC 178629

|||||

QY 555

GCCTTGGGGCCACTAGCGGGCGCTCTCCAGTGCAGCTCTTCCCTTGCCTTC 614

|||||

Db 178630

GCCTTGGGGCCACTAGCGGGCGCTCTCCAGTGCAGCTCTTCCCTTGCCTTC 178689

|||||

QY 615

ACTGTGGAGCGCGGACCTTGACCCGAGGGGGCGCGCGCGCTGGGAAGTCTT 674

|||||

Db 178690

CTGTGTGGAGCGCGGACCTTGACCCGAGGGGGCGCGCGCTGGGAAGTCTT 178749

|||||

QY 675

CGACGTGTGGAGCGGCTGCGCCAGCCAGCCCTGGAAGAGCTGTGTGGAGCT 734

|||||

Db 178750

CGACGTGTGGAGCGGCTGCGCCAGCCAGCCCTGGAAGAGCTGTGTGGAGCT 178809

|||||

QY 735

CGCATGGGGGAGCTGAGCGCGGGGAGGCGCGCGGGGACCCAGCAACC 794

|||||

Db 178810

CGCATGGGGGAGCTGAGCGCGGGGAGGCGCGCGGGGACCCAGCAACC 178869

|||||

QY 795

GCCGCGCGGAGCTGCGGAGTCTGGCTTTCGCGGAGGGTGGCGCTCCCGAGG 854

|||||

Db 178870

GCGCGCGGAGCTGCGGAGTCTGGCTTTCGCGGAGGGTGGCGCTCCCGAGG 178929

|||||

QY 855

GGCCCTGCTGGTGTATTACACAGATCCAGCGCAAGAACTTTCGACAGATGCG 914

|||||

Db 178930

GGCCCTGCTGGTGTATTACACAGATCCAGCGCAAGAACTTTCGACAGATGCG 178989

|||||

QY 915

GCAGCTGGGCTCGCGAGGCTCGCGGGCGCGGGGGCGCGAGGGTCTGTGGCG 974

|||||

Db 178990

GCAGCTGGGCTCGCGAGGCTCGCGGGCGCGGGGGCGCGAGGGTCTGTGGCG 179049

|||||

QY 975

GCGCTGGGGCGCGCGAGGCTTGGCTGCGCTGCGCGCGCGCGCGCGCGCGCG 1034

|||||

Db 179050

GCGCTGGGGCGCGCGAGGCTTGGCTGCGCTGCGCGCGCGCGCGCGCGCGCG 179109

|||||

QY 1035

CACGGCTTTCGCCAGTCCGATGCCAAGCGGCACGCAAGAGTCCAGGCTGCTG 1094

```
|||||
Db 179110 CACGGCTTCGCGAGCGCCATGGAAGCGTCACGCAAGAGTCCAGGCTACGCTGCAG 179169
QY 1095 CAAGAAGCCCTGCAGTGAACCTCAAGAGAGCTGGCTGGGAGCACTGATTATCGCGCC 1154
Db 179170 CAAGAAGCCCTTCGACGTGAACCTCAAGAGAGCTGGCTGGGAGCACTGATTATCGCGCC 179229
QY 1155 COTGGAGTACGAGGCTATCACTGGGAGGCTATGCGACTTCCGCTCCGCTCGCACCT 1214
Db 179230 COTGGAGTACGAGGCTATCACTGGGAGGCTATGCGACTTCCGCTCGCACCT 179289
QY 1215 GGAGCCCAACCAACGCGCATCATCAGACGCTGATGAACCTCCATGGACCCCGGCTCCAC 1274
Db 179290 GGAGCCCAACCAACGCGCATCATCAGACGCTGATGAACCTCCATGGATCCCGGCTCTAC 179349
QY 1275 CCGGCCAGCTGCTGGTGGCCCAACCAATTGACTCCCATCAGCATCTATATCAGACGC 1334
Db 179350 TCGGCCAGCTGCTGGTGGCCCAACCAATTGACTCCCATCAGCATCTGATATCAGACGC 179409
QY 1335 GGGCAATATGTTGCTACAACGACTACGAGGACATGGTGGAGTGTGGGCTGCAG 1394
Db 179410 GGGCAATATGTTGCTACAACGACTACGAGGACATGGTGGAGTGTGGGCTGCAG 179469
QY 1395 GTAGCGGTGCTTTCCCGCGCGCTTGGCGCG 1425
Db 179470 GTAGCGGTGCTTTCCCACTGCTTGGCGCG 179500

RESULT 8
BTU13661
LOCUS
DEFINITION Bos taurus cartilage-derived morphogenetic protein 2 (CDMP-2) mRNA,
complete cds.
ACCESSION U13661
VERSION U13661.1 GI:632489
KEYWORDS
SOURCE
ORGANISM Bos taurus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 1308)
Chang,S.C., Hoang,B., Thomas,J.T., Vukicevic,S., Luyten,F.P.,
Ryba,N.J., Kozak,C.A., Reddi,A.H. and Moos,M. Jr.
Cartilage-derived morphogenetic proteins. New members of the
transforming growth factor-beta superfamily predominantly expressed
in long bones during human embryonic development
J. Biol. Chem. 269 (45), 28227-28234 (1994)
JOURNAL 95050604
MEDLINE 7961761
REFERENCE 2 (bases 1 to 1308)
Moos,M.
Direct Submission
Submitted (18-AUG-1994) Malcolm Moos, FDA/CBER, 1401 Rockville Pike
(HFM-527), Rockville, MD 20852-1448, USA
On Jan 24, 1995 this sequence version replaced gi:600733.
COMMENT Location/Qualifiers
FEATURES
1..1308
/organism="Bos taurus"
/db_xref="taxon:9913"
/tissue_type="articular cartilage"
/dev_stage="adolescent"
1..1308
/gene="CDMP-2"
<1..>1308
/gene="CDMP-2"
/codon_start=1
/product="cartilage-derived morphogenetic protein 2
precursor"
/protein_id="AA61416.1"
/db_xref="GI:632490"
/translation="RASAELGSAKGMTREKGRMPRAPRENATAREPLDRQPPRPQ
EPPQRRPQPEAREPPGRGRLVPHEYMLSIYRTYIAEKLGINASFFQSSKSANTI
```

```
TSFVDRGLDLSHTPLRRQKYLFDVSTLSDKEELVGADVRLFRQAPALAPPAAPLA
ALRUPVAPAGSAPGAPGAVPQWVDFVNRGLRPOPWKLCLELRAWGEPCAAE
DEARTPGPOPPDLRSLGFRVRVTPQERALLVFSRSORKTTLFAMRQLGSATE
VVGCGGAGSGPPPPPPPPSGTDPAGLWSPSPRRRTAFASRHKRKGKSLRL
CSKRLPHVNFKELGWDWIIAPLEYATHCGVCDFFPLRSHLEPTNHAIITLMSMD
PGSTPPSCCVPTKLTPISILYIDAGNNVYVNEYEMVYVSCGR"
949..1308
/mat_peptide
/gene="CDMP-2"
/misc_feature
265..273
/product="cartilage-derived morphogenetic protein 2"
/gene="CDMP-2"
/misc_feature
937..948
/note="encodes putative N-linked glycosylation site"
/gene="CDMP-2"
/note="encodes putative RXR proteolysis site"
BASE COUNT 209 a 456 c 452 g 191 t
ORIGIN
Query Match 65.1%; Score 928.2; DB 4; Length 1308;
Best Local Similarity 86.5%; Pred. No. 7.3e-137;
Matches 1136; Conservative 0; Mismatches 123; Indels 54; Gaps 8;
QY 128 COTCCGCGAGCTGGTTCACCAAGGGCATCGAAGCGCGCAAGCAAGATGCGAC 187
Db 5 CGTCCGCGAGCTGGCTCCGCCAAGGGAATCGAAGCGCGCAAGGAAGATGCCGC 64
QY 188 GGGCGCGCGGACAGTACGCGGGCGCGGAG-----GGCCAGGAACACAGCGCG 238
Db 65 GGGCGCGGAGAGAAATGCCACGCCGGGAGCCCTGGATCGCCAGGAGCCCCCGGA 124
QY 239 GGCCTCAGGACGAACCC-----CGGGCTCAGCAGCCCGCGGCGAGGAGCGCCAG 289
Db 125 GGGCGCGGAGGAGCCCGCGCGCGCCACACAGCAGCTCGGAGCGCTCCCG 184
QY 290 GCAGGGGTCGCGCGCTGGTGCCCGACAGATACATGCTGTCAATCTACAGGACTTACTCA 349
Db 185 GCAGGGGCGCGCTGGTGCCCGACAGATACATGCTGTCAATCTACAGGACTTACTCA 244
QY 350 TCGCTGAGAAGCTGGGATCAATGCCAGCTTTTCCAGTCTTCCAGTGGCTTAACGA 409
Db 245 TCGCCGAGAAGCTGGGATCAATGCTTTTCCAGTCTTCCAGTGGCTTAACGA 304
QY 410 TCACAGCTTTGTAGCAGGGGACTAGAGTCTCTCGCACACTCTCTCCGGAGACAGA 469
Db 305 TCACCTAGCTTTGTAGCAGGGGACTAGAGTCTCTCGCACACTCTCTCCGGAGACAGA 364
QY 470 AGTATTGTTGATGTCTCATGTCTCAGACAAAGAGAGCTGTGGCGCGGAGCTGC 529
Db 365 AGTATTGTTGATGTCTCATGTCTCAGACAAAGAGAGCTGTGGCGCGGAGCTGC 424
QY 530 GGCCTTTTCGCGAGCGCCCTCAGCGCCCTGGGGGCCACACAGCGCGGCTCCACGTGC 589
Db 425 GGCCTTTTCGCGAGCGCCCTGCTGCGCTGGCGCCCGCCG-----GGCGCTCCGCTGC 479
QY 590 AGCTCTCCCTTTCCTTTTCGCGCTTCTGCTGGAGCGCGGACCCCTGGACCGCGAGGGG 649
Db 480 AGCTCTCCCTTTCCTTTCGCGCTTCTGCTGGAGCGCGGAGCGCTGGACCGCA-GGG 538
QY 650 CGCCCGCGCGCGCTGGGAAGTCTTCGAGGTGTGCGAGGCGCTGCGCCACACAGCCCTGA 709
Db 539 CGCCCGCGCGCGCTGGGAAGTCTTCGAGGTGTGCGAGGCGCTGCGCCCGCGCGCTGGA 598
QY 710 AGCAGCTGTGCTTGGAGCTGGGGCGCGCATGG---GGCAGCTGCACGCGGGAGGCGG 766
Db 599 AGCAGCTGTGCTTGGAGCTTGGGGCGCGGTGGGGCGCGGAGCGCGCGCGCGGAGGAG 658
QY 767 AGGCGCGCGCGGGGACCCAGCAACCGCGCCCGCCCGGACCTCGCGAGTCTGGGCTTCG 826
Db 659 AGGCGCGCACGCTTGGGCGCGCGAGCGCGCGCCCGCGGAGCTCGGAGTCTGGGCTTCG 718
QY 827 GCGGAGGCTGGGCGCTCCCGAGGAGCGGCGCTCTGCTGGTGTATTACAGCATCCGACG 886
Db 719 GCGGAGGCTGGGAGCGCGCGCGGAGCGCGCTTGTCTGCTGTGTCTTCCAGGTCGCCAGC 778
```



QY 1283 GCTGCTGGTGGCCACCAAAATGACTCCCATCAGCATCTATACATCAGCGGGCAATA 1342  
|||||  
Db 1196 GCTGCTGGTGGCCACCAAAATGACTCCCATCAGCATCTGTATACATCAGCGGGCAATA 1255  
|||||  
QY 1343 ATGTGGTCTACAAGACGATCAGGAGGAGTGGTGGTGGAGTGGTGGCGGCTGCAGG 1395  
|||||  
Db 1256 ATGTGGTCTACAAGACGATCAGGAGGAGTGGTGGTGGAGTGGTGGCGGCTGCAGG 1308  
|||||

RESULT 10  
AC126238 143312 bp DNA linear HTG 04-JUL-2002  
LOCUS  
DEFINITION  
Felis catus clone RP86-410H13, WORKING DRAFT SEQUENCE, 7 unordered  
pieces  
ACCESSION  
AC126238  
VERSION  
AC126238.1 GI:21693902  
KEYWORDS  
HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE  
cat.  
ORGANISM  
Felis catus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
1 (bases 1 to 143312)  
Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,  
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,  
Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J.,  
Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Laric,P.,  
Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,  
Margulies,E.H., McDowell,J., Maskeri,B., Mastrian,S.D.,  
McCluskey,J.C., McEwell,J., Paguirigan,C., Pearson,R.,  
Portnoy,M.E., Prasad,A., Schueler,M.G., Stantripop,S., Thomas,J.W.,  
Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A.,  
Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.  
NISC Comparative Sequencing Initiative  
Unpublished  
2 (bases 1 to 143312)  
Green,E.D.  
Direct Submission  
Submitted (04-JUL-2002) NIH Intramural Sequencing Center, 8717  
Groveomont Circle, Gaithersburg, MD 20877, USA  
----- Genome Center  
Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: http://www.nisc.nih.gov  
Contact: nisc.zoo@nih.gov  
----- Project Information  
Center project name: czs  
Center clone name: 410H13  
----- Summary Statistics  
Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 140814 bases at least Q40  
Consensus quality: 141228 bases at least Q30  
Consensus quality: 141484 bases at least Q20  
Insert size: 142000; agarose-fp  
Insert size: 142712; sum-of-contigs  
Quality coverage: 9.96x in Q20 bases; agarose-fp  
Quality coverage: 9.91x in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 7 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 2817: contig of 2817 bp in length  
\* 2818 2917: gap of unknown length  
\* 2918 17821: contig of 14904 bp in length  
\* 17821 17921: gap of unknown length  
\* 17922 37586: contig of 19665 bp in length

37587 37686: gap of unknown length  
\* 37687 56500: contig of 18814 bp in length  
\* 56501 56600: gap of unknown length  
\* 56601 73714: contig of 19114 bp in length  
\* 73715 75814: gap of unknown length  
\* 75815 109353: contig of 33539 bp in length  
\* 109354 109453: gap of unknown length  
\* 109454 143312: contig of 33859 bp in length.  
FEATURES  
Location/Qualifiers  
1..143312  
/organism="Felis catus"  
/db\_xref="taxon:9685"  
/clone="RP86-410H13"  
/clone\_lib="RP86"  
1..2817  
/note="assembly\_fragment"  
2918..17821  
/note="assembly\_fragment"  
17922..37586  
/note="assembly\_fragment"  
37687..56500  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:right  
56601..75714  
/note="assembly\_fragment"  
75815..109353  
/note="assembly\_fragment"  
109454..143312  
/note="assembly\_fragment"  
BASE COUNT 39933 a 31312 c 30631 g 40835 t 601 others  
ORIGIN  
Query Match 62.2%; Score 885.8; DB 2; Length 143312;  
Best Local Similarity 94.0%; Pred. No. 1.6e-130;  
Matches 933; Conservative 0; Mismatches 57; Indels 3; Gaps 1;  
QY 435 AGACGATCTCTCGCACACTCTCTCCGGAGACAGAAATATTTTGTGATGTGTCATGCT 494  
|||||  
Db 50385 AGACGATCTCTCGCACACTCTCTCCGGAGACAGAAATATTTTGTGATGTGTCACACT 50444  
|||||  
QY 495 CTCAGACAAAGAGAGCTGGTGGCGCGAGCTCGGCTCTTTTCGCCAGGCCCTCAGC 554  
|||||  
Db 50445 CTCAGACAAAGAGAGCTGGTGGCGCGAGCTCGGCTCTTTTCGCCAGGCCCTCAGC 50504  
|||||  
QY 555 GCCCTGGGGGCCACAGCGGGCCGCTCCACGTTCAGCTCTTCCTTGCCTTTCGCCCT 614  
|||||  
Db 50505 GCCCTGGGGGCCGCGCGGCCGCTCCACGTTCAGCTCTTTCCCTGCTGCGCCCA 50564  
|||||  
QY 615 ACTGCTGGACGCGCGGACCTTGGACCCGCGAGGGCGCGCGCGGCTGGGAAGTCTT 674  
|||||  
Db 50565 GCTGCTGGACGCGCGGACCTTGGACCCGCGAGGGCGCGCGCGGCTGGGAAGTCTT 50624  
|||||  
QY 675 CGACGTGTGGCAGGGCCCTGCGCCACACAGCCCTGGAAGCAGCTGTGCTTGGAGCTGCGGC 734  
|||||  
Db 50625 CGACGTGTGGCAGGGCCCTGCGCCACACAGCCCTGGAAGCAGCTGTGCTTGGAGCTGCGGC 50684  
|||||  
QY 735 CGCATGGGGCAGCTGGAGCGCGGGAGCGCGCGCGCGCGCGGGGACC---CCAGCA 791  
|||||  
Db 50685 CGCGTGGGGCAGCGCGGGAGCTGGGGAGCGCGAGCACGCGCGCGCGCGCCAGCA 50744  
|||||  
QY 792 ACCGCGCGCGCGGACCTTGGGAGTCTGGGCTTTCGCGGAGGGTTCGCGCTCCCGCAGGA 851  
|||||  
Db 50745 GCCACCGCGCGGACCTTGGGAGTCTGGGCTTTCGCGGAGGGTTCGCGCGCGCGCGCCAGGA 50804  
|||||  
QY 852 GCGGGCCCTGCTGTGGTATTACACAGATCCACAGCGCAAGAACCTGTTCGCGAGATGCG 911  
|||||  
Db 50805 GCGGCGCCCTGCTGTGCTGTTTCCACAGATCCACAGCGCAAGAACCTGTTCGCGAGATGCG 50864  
|||||  
QY 912 CGACGAGCTGGGCTCGGCGCGGAGGCTGCGGGCGCGCGCGCGCGGGGCTGCTGCGCC 971  
|||||  
Db 50865 CGACGAGCTGGGCTCGGCGCGGAGGCTGCGGGCGCGCGCGCGCGGGGCTGCTGCGCC 50924  
|||||



QY	972	GCCGCGCTCGGGCCCGCCGAGGCTTGGCTGCCCTCCCGCGCGCGCGCGGCG	1031
Db	50925		
QY	1032	GCCGAGGCTTCGCCAGTCGCATGGCAGCGGCACGCAAGAGTCAGGCTAGCGTG	1091
Db	50985		
QY	1092	CAGCAAGAAGCCCTCAGCTGAACCTTCAAGGAGCTGGGCTGGGAGCACTGGATTATCGC	1151
Db	51045		
QY	1152	GCCCTGGAGTAGAGGCTTATCAGTCGAGGGTGTATGGACTTCCGCTGCGCTCGCA	1211
Db	51105		
QY	1212	CTGTGAGCCCAACACAGCCATCATCAGAGCGTGTATGAACTCCATGATGATCCCGGCTC	1271
Db	51165		
QY	1272	CACCCGCGCCAGCTGTGCTGCCGCCACCAATTTACTCCCATCAGCATTTATACATCGA	1331
Db	51225		
QY	1332	CGCGGCAATATGTGCTTACAGCAGTACGAGGACATGGTGGTGGAGTCTGCGGCTG	1391
Db	51285		
QY	1392	CAGGTAGCGGTGCTTTCCCGCGCTTGCGCC	1424
Db	51345		
LOCUS	AC126232/c		
DEFINITION	Bos taurus clone RP42-416E22, WORKING DRAFT SEQUENCE, 12 unordered		
ACCESSION	AC126232		
VERSION	AC126232.1 GI:21693896		
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.		
SOURCE	cow.		
ORGANISM	Bos taurus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Butheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
	Bovidae; Bovinae; Bos.		
	1 (bases 1 to 219978)		
	Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,		
	Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,		
	Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J.,		
	Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Laric,P.,		
	Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,		
	Margulies,E.H., Mastello,C., Maskeri,B., Mastrian,S.D.,		
	McCloskey,J.C., McDowell,J., Paguirigan,C., Pearson,R.,		
	Porcnoy,M.E., Prasad,A., Schueler,M.G., Stantripop,S., Thomas,J.W.,		
	Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.		
TITLE	NISC Comparative Sequencing Initiative		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 219978)		
AUTHORS	Green,E.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (04-JUL-2002) NIH Intramural Sequencing Center, 8717		
COMMENT	Grosvont Circle, Gaithersburg, MD 20877, USA		
	----- Genome Center		
	Center: NIH Intramural Sequencing Center		
	Center code: NISC		
	Web site: <a href="http://www.nisc.nih.gov">http://www.nisc.nih.gov</a>		
	Contact: nisc_zoehghri.nih.gov		
	----- Project Information		
	Center project name: djb		
	Center clone name: 416E22		
	----- Summary Statistics		

Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 213653 bases at least Q40  
Consensus quality: 214882 bases at least Q30  
Consensus quality: 215481 bases at least Q20  
Insert size: 209000; agarose-fp  
Insert size: 218878; sum-of-contigs  
Quality coverage: 7.79x in Q20 bases; agarose-fp  
Quality coverage: 7.44x in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence.  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*  
\* 1 3801: contig of 3801 bp in length  
\* 3802 3901: gap of unknown length  
\* 3902 6418: contig of 2517 bp in length  
\* 6419 6518: gap of unknown length  
\* 6519 12134: contig of 5616 bp in length  
\* 12135 12234: gap of unknown length  
\* 12235 18070: contig of 5836 bp in length  
\* 18071 18170: gap of unknown length  
\* 18171 37754: contig of 19584 bp in length  
\* 37755 37854: gap of unknown length  
\* 37855 57704: contig of 19850 bp in length  
\* 57705 57805: gap of unknown length  
\* 57806 84717: contig of 26913 bp in length  
\* 84718 108293: contig of unknown length  
\* 108294 108393: gap of unknown length  
\* 108394 130107: contig of 21714 bp in length  
\* 130108 130207: gap of unknown length  
\* 130208 155272: contig of 25065 bp in length  
\* 155273 183289: contig of unknown length  
\* 183290 183389: gap of unknown length  
\* 183390 219978: contig of 36589 bp in length.  
\*  
\* Location/Qualifiers  
\* 1. 219978  
\* /organism="Bos taurus"  
\* /db\_xref="taxon:9913"  
\* /clone="RP42-416E22"  
\* /clone\_lib="RP42"  
\* 1. 3801  
\* /note="assembly\_fragment"  
\* 3902. 6418  
\* /note="assembly\_fragment"  
\* 6519. 12134  
\* /note="assembly\_fragment"  
\* 12235. 18070  
\* /note="assembly\_fragment"  
\* 18171. 37754  
\* /note="assembly\_fragment"  
\* 37855. 57704  
\* /note="assembly\_fragment"  
\* clone\_end:SP6  
\* vector\_side:right  
\* 57805. 84717  
\* /note="assembly\_fragment"  
\* 84818. 108293  
\* /note="assembly\_fragment"  
\* 108394. 130107  
\* /note="assembly\_fragment"  
\* 130208. 155272  
\* /note="assembly\_fragment"  
\* 155373. 183289  
\* /note="assembly\_fragment"  
\* 183390. 219978

```
/note="assembly fragment"
```

BASE COUNT	59915 a	49896 c	49290 g	59765 t	1112 others
ORIGIN	/note assembly fragment				

Query Match	56.1%	Score 798.8	DB 2	Length 219978
Best Local Similarity	89.7%	Pred. No. 7.7e-117		
Matches 913	Conservative 0	Mismatches 177	Indels 28	Gaps 4
QY 435	AGACGATCTCTCGCACACTCTCTCCGGAGACAGAGCTATTGTGTGATGTGTCCATGCT	494		
DB 126101	AGACGATCTCTCGCACACTCTCTCCGGAGACAGAGTATTGTGTGATGTGTCCACGCT	126042		
QY 495	CTCAGACAAAGAGAGTGTGGCGGGAGCTCGGGCTCTTTGCCAGGCGCCCTCAGC	554		
DB 126041	CTCAGACAAAGAGAGTGTGGCGGGAGCTCGGGCTGTTCGCCAGGCGCCGCTGC	125982		
QY 555	GCCCTGGGGGCCACAGCGGGGCCGTCCACGTGAGCTCTTCCCTTGCCCTTTCCGCCCT	614		
DB 125981	GCCCTGGGGGCCGCGCGCGCGCTCCGCTTTCAGCTCTTCCGCTTCCGCTGCCAGTGC	125922		
QY 615	ACTCTGAGCGCGGACCTTGAGCCGCGGGGGCGCGCGCGCGCTGGGAAGTCTT	674		
DB 125921	GCTGCTGGAACGCGGAGCTTGAGCCGAGGGGGCGCCCGCGCGCTGGGAAGTCTT	125862		
QY 675	CGACGTGTGCGAGGGCTTGCGCCACCAACCCCTGAAAGCAGCTGTCTTGGASCCTGCGGC	734		
DB 125861	CGACGTGTGCGGGGCGCTTGCGCCCGCCAGCCCTGGAAGCAGCTGTCTTGGAGCTTCGGGC	125802		
QY 735	GGCATGG---GGCGAGCTGAGCGCGGGGAGCGAGCGCGCGCGGGGACCCACGCA	791		
DB 125801	CGCTGTGGGGCGGAGCGGGCGCGGAGGAGCGAGCGCGCGCGCTGGGCCCCACGCA	125742		
QY 792	ACCGCCGCCCGGACCTCGGAGCTCTGGGCTTCGGCGCGGAGGTGGGGCTCCGCCAGGA	851		
DB 125741	GCCGCCGCCCGGACCTCGGAGTCTGGGCTTCGGCAGGAGGTGCGGACCCCGCCAGGA	125682		
QY 852	CGGGGCCCTGCTGTGTATTCACAGATCCACGCGCAAGAACCTGTTTCGCAGAGATGCG	911		
DB 125681	CGCGGCCCTGCTGTGTGTCTCCAGGTCCACGCGAAGAACCTGTTTCGCCGAGATGCG	125622		
QY 912	CGAGCAGCTGGGCTCGG---CCGAGGCTCGGGCGCGGGCGCGCGGAGGGTCTGT	968		
DB 125621	CGAGCAGCTGGGCTCGGCGACCGAGGTGTTCGGCCCGGTGTGGGGCCGAGGGTCTGG	125562		
QY 969	-----GCGCGCGCGTCTGGGGCGCCCGGATGCCAGGCGCTTGCT	1007		
DB 125561	GCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGGGCTCTGGTC	125502		
QY 1008	GCCCTCGCGCGCGCGCGGGCGCGCGCGCTTCGCCAGTGCCTATGCCAGCGGCA	1067		
DB 125501	GCCCTCGCGCTGGCGCGGGCGGGCGCGCGCGCTTCGCCAGCGCGCGCAAGCGGCA	125442		
QY 1068	CGGCAAGAAGTCCAGGTACGCTCGACAGAGCCCTGCACCTGAACTTCAAGGAGCT	1127		
DB 125441	CGGCAAGAAGTCCAGGCTGGCTCGACAGAGCCCTGCACCTGAACTTCAAGGAGCT	125382		
QY 1128	GGGCTGGGAGCTGGATTATTCGGCGCCCTGGAGTACAGGCGCTATCACTCGGAGGTGT	1187		
DB 125381	GGGCTGGGAGCTGGATTATTCGGCGCCCTGGAGTACAGGCGCTTACCAGTTCGAGGCGGT	125322		
QY 1188	ATGGGACTTCCGCTGCGCTCGCACTTGGAGCCCAACCAACCGCCCATCTCCAGAGCGCT	1247		
DB 125321	GTGGGACTTCCGCTAGCTCGCACTTGGAGCCCAACCAACCGCCCATCTCCAGAGCGCT	125262		
QY 1248	GATGAATCCATGAGCCCCGGCTCCACCGCGCCAGTGTGTGCTGCGCCCAAAATTGAC	1307		
DB 125261	GATGAATCCATGAGCCCCGGCTCCACCGCGCCAGTGTGTGCTGCGCCCAAAATTGAC	125202		
QY 1308	TCCATCAGACTTATACATCGAGCGGGCAATAATGTGCTACAGCAGTACGAGGA	1367		
DB 125201	TCCATCAGACTTGTACATCGAGCGGGCAATAATGTGCTACAGCAGTACGAGGA	125142		
QY 1368	CATGGTGGTGGAGTCGTGCGGCTCGAGGTACGGGTGCTTTTCCGCGCGCGCTTGCCCG	1425		

Db 125141 GATGGTGGTGGAGTCGTGCGGCTGACAGGTAGCGGCGCC-TTCTGCGCGCTCGGCCCCG

RESULT 12  
AC058786/C

LOCUS	AC058786	209733 bp	DNA	linear	ROD 29-JAN-2002
DEFINITION	Mus musculus clone rp23-1107, complete sequence.				
ACCESSION	AC058786				
VERSION	AC058786.23	GI:11120825			
KEYWORDS	HTG.				

SOURCE ORGANISM

mus musculus  
unpublished  
2 (bases 1 to 209733)  
Pan, H. and Roe, B. A.  
Direct Submission  
JOURNAL  
Submitted (19-APR-2000) Department Of Chemistry And Biochemistry.

## REFERENCE

3 (Jases 1 CO 209733)  
AUTHORS  
Pan, H. and Roe, B.A.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (08-NOV-2000) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA

## REFERENCE

**AUTHORS** Pan, H. and Roe, B. A.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (29-JAN-2002) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA

**COMMENT**

----- Genome Center  
Center: Department Of Chemistry And Biochemistry  
The University Of Oklahoma  
Center code:UOKNOR

## FEATURES

```

1. .209733
   source
   accession=RP23
   organism="Mus musculus"
   db_xref="taxon:10090"
   clone="rp23-117o7"
   clone_lib="RP23"

```

BASE COUNT	61917 a	45381 c	44340 g	58095 t
ORIGIN				

Query Match	53.7%	Score 765.8;	DB 10;	Length 209733;
Best Local Similarity	87.4%	Pred. No. 1.2e-111;		
Matches 866;	Conservative	0;	Mismatches 112;	Indels 13;
				Gaps 2;

**QY**      435 AGACGATCTCTCGCACACTCCTCTCCGGAGACAGAAGTATTGTTTTGATGTGTCCATGCT 494  
|||||  
65214 |||||  
**Db**      65214 AGACGATCTCTCGCACACTCCTCTCCGGAGACAGAAGTATTGTTTTGATGTGTCCACACT 65155  
|||||

[illegible]

Qy	555	GCCCTGGGGGCCACGACGGGGCCGCTCCAGCTGCAGCTCTCCCTTGCCTTTCGCCCT	614
Db	65094	GCCTTGGGGGCTACGGCCGACCGCTGCAGCTCTCCCTTGCCTTTCGCCCT	65035

Qy	615	ACTGCTGGAGCGCGGACCTTGGACCCGCGAGGGGGCGCGCGCGCTGGGAAGCTCTT	674
Db	65034	GCTACTGGAGSCCAGGACCTCTGGATCTCAGGGACCAACCCAGGCCGCTGGGAAGCTCTT	64975

QV 675 CGACGTGTGGCAGGGCCCTGCGCCACCCAGCCCTGGAAGCAGCTGTGCTTGGAGCTGCGGGC 734



[illegible]

\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 4963: contig of 4963 bp in length  
\* 4964 5063: gap of unknown length  
\* 5064 38133: contig of 33070 bp in length  
\* 38134 38233: gap of unknown length  
\* 38234 91212: contig of 52979 bp in length  
\* 91213 91312: gap of unknown length  
\* 91313 167312: contig of 76000 bp in length.

## FEATURES

source  
1. .167312  
/organism="Rattus norvegicus"  
/strain="Brown Norway"  
/db\_xref="taxon:10116"  
/clone\_lib="RP31"  
/clone\_lib="RP31"  
1. .4963  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:right  
5064. .38133  
/note="assembly\_fragment"  
clone\_end:T7  
vector\_side:right  
38234. .91212  
/note="assembly\_fragment"  
91313. .167312  
/note="assembly\_fragment"  
BASE COUNT 47363 a 35040 c 36627 g 47981 t 301 others  
ORIGIN

Query Match 52.6%; Score 749.8; DB 2; Length 167312;  
Best Local Similarity 86.4%; Pred. No. 4.2e-109;  
Matches 856; Conservative 0; Mismatches 122; Indels 13; Gaps 2;

QY 435 AGAGATCTTCGCACACTCTCTCCGACACAGAGTATTGTTGATGTGCCATGCT 494  
|||||  
Db 76193 AGAGATCTTCGCACACTCTCTCCGACACAGAGTATTGTTGATGTGCCATGCT 76252  
|||||

QY 495 CTCAGACAAAGAGAGTGTGGCGCGGAGCTGCGGCTCTTTGCCAGCGCCCTCAGC 554  
|||||  
Db 76253 CTCAGACAAAGAGAGTGTGGCGCGGAGCTGAGGCTTTATCGCCAGCGCCCAAC 76312  
|||||

QY 555 GCCTGGGGCCACCAGCCGGCGCTCCAGCTGCAGCTCTTCCCTTGCCCTTTCGCCCT 614  
|||||  
Db 76313 GCCTGGGGCCCAACACCAGCGCTGCATTTGCAGCTCTTCCCTGTTTGTCCCAT 76372  
|||||

QY 615 ACTGCTGGAGCGCGGACCTTGACCGCGAGGGGCGCGCGCGGCTGGGAAGTCTT 674  
|||||  
Db 76373 GTTACTGGACTCTAGGACCTTGATCTCAGGGGCCAACAGAGCTGGCTGGGAAGTTT 76432  
|||||

QY 675 CGAGCTGTGCAGGGCTGCGCCACCAGCCCTGGAAGCAGCTGTGCTGGAGCTGCGGGC 734  
|||||  
Db 76433 CGAGCTGTGCAGTCTCGGCCCTCAGCCGTGGAAGCAGCTGTGCTGGAGCTTGGGC 76492  
|||||

QY 735 CGATGGGGCGAGCTGGACCGCGGGAGCGGAGGCGCGCGGGGACCCAGCAACC 794  
|||||  
Db 76493 AGCTGGGGTGAGCTGGACCGCCAGGGATTACAGGAGCGCGCGGGGTCCTCCCAACAGTC 76552  
|||||

QY 795 GCGGCCCGGACCTGGGAGTCTGGGCTTCGCGCGGAGGCTGGGCTCTCCAGGAGCG 854  
|||||  
Db 76553 ACCGCCCTGGAGCTGGGAGTCTGGGCTTCGCGCGGAGGCTGGGCGCGCCAGGAGCG 76612  
|||||

QY 855 GGCCCTGCTGGTGGTATTCACAGATCCACAGCGCAAGAACTGTTCGACAGATGCGCGA 914  
|||||  
Db 76613 CGCCCTGCTGTGTTTACAGATGCGAGCGCAAGAACTGTTCACATGAGATGATGA 76672  
|||||

QY 915 GCAGCTGGCTCGGCGAGAGCTCGGGGCCCGGGCGCGGGGCTGTGTGGCGCGC 974  
|||||

Db 76673 GCAGCTGGGGCTCTGCAGAGGCT-----GCCGGAGCCGAGGGGTCTATGGCCAGC 76720  
QY 975 GCCTGCGGGCCCGCGGATCCAGAGCCCTTGCTGCCCTCCCGCGCGCGCGCGCGCGCG 1034  
|||||  
Db 76721 GCCGTGGGGCCCGCGGAGCGCGGGTCTTGCTGCCCTCCCGCGCGCGCGCGCGCGCG 76780  
|||||

QY 1035 CAGGCGCTTCGCCAGTGCCTATGCGGCAAGCGCGGCAAGTCCAGGCTACGCTCGAC 1094  
|||||  
Db 76781 CACCGCCCTCAGCAGCGCTCACGCAAGCGACACGCGCAAGTCCAGGCTGCGCTCGAC 76840  
|||||

QY 1095 CAAGAGCCCTTCAGCTGACCTTCAGGAGCTGGGCTGGAGCAGCTGGATTATCGCGCC 1154  
|||||  
Db 76841 CAGAAACCTCTGCAGCTGAATTTTAAAGAGTAGGCTGGGAGCATTTGATTCGCGCC 76900  
|||||

QY 1155 CTTGGAGTACGAGGCTTATCACTGCGAGGCTGTATGCACTTCCCGCTGCGCTCGACCT 1214  
|||||  
Db 76901 CTTAGAGTACGAGGCTTATCACTGCGAGGCGTCTGCGACTTTCCGCTGCGCTCGACCT 76960  
|||||

QY 1215 GGAGCCCAACACACGCCATCATCCAGAGCTGATGAATCCATGAGACCCCGGCTCCAC 1274  
|||||  
Db 76961 GGAGCCCAACCAACCATGCCATCATCCAAACGCTGATGAATCCATGAGACCCCGGCTCCAC 77020  
|||||

QY 1275 CCGCGCCAGCTGCTGCCGCCACCAATTTGACTCCCATCAGCATTTCTATACATCGAGC 1334  
|||||  
Db 77021 CCCACCCAGCTGCTGTTCACCAAAATTTGACTCCCATCAGCATTTCTATACATCGATGC 77080  
|||||

QY 1335 GGGCAATAATGTGCTACAAAGCAGTACAGAGCATGGTGGTGGAGTCGCGGCTCGAC 1394  
|||||  
Db 77081 GGGCAATAATGTGCTACAAAGCAGTACAGAGCATGGTGGTGGAGTCGCTGTGGTTGTAG 77140  
|||||

QY 1395 GTAGCGGTGCTTTTCGCGCGCGCTTGGCCCG 1425  
|||||  
Db 77141 GTAGCGGTG-CTGTCCCGCTCTCTGGGCCAG 77170  
|||||

RESULT 15  
LOCUS 162382  
DEFINITION Sequence 31 from patent US 5658882.  
ACCESSION 162382  
VERSION 162382.1 GI:2480330  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1046)  
AUTHORS Celeste,A.J., Wozney,J.M., Rosen,V.A., Wolfman,N.M., Thomsen,G.H.  
and Melton,D.A.  
TITLE Methods of inducing formation of tendon and/or ligament tissue  
comprising administering BMP-12, BMP-13, and/or MP-52  
JOURNAL Patent: US 5658882-A 31 19-AUG-1997;  
FEATURES Location/Qualifiers  
Source 1. .1046  
BASE COUNT 212 a 314 c 331 g 187 t 2 others  
ORIGIN

Query Match 40.8%; Score 581.8; DB 6; Length 1046;  
Best Local Similarity 87.2%; Pred. No. 2.8e-82;  
Matches 675; Conservative 1; Mismatches 84; Indels 14; Gaps 3;

QY 653 CGCGCGCGCGCTGGGAAGTC-TTCGACGTGTGGCAGGCGCTTGGGCCACACAGCCCTTGAAG 711  
|||||  
Db 59 CCCAGCGCGCTGGGAAGTCTTTTCGAGCTGTGGCAGGCGCTTGGCCCTCAGCCCTTGAAG 118  
|||||

QY 712 CAGCTGTGCTTGGAGCTGCGGGCGCGCATGCGGAGAGTGCAGCGCGGGAGCGAGCGG 771  
|||||  
Db 119 CAGCTGTGCTTGGAGTGTGGCGGAGCGCTGGGGTGAAGTGCAGTGCAGTGCAGTGCAG 178  
|||||

QY 772 CGCGCGGGGACCCAGCAACCGCGCCCGCGGAGCTGCGGAGCTCTGGGCTTCGCGCGG 831  
|||||  
Db 179 CGCGCGAGGGGTCCCCAGCAGCAACCGCCCTTGAGACCTGCGGAGTCTTGGGCTTCGGGTCG 238  
|||||



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 01:01:20 ; Search time 78 Seconds  
(without alignments)  
5602.749 Million cell updates/sec

Title: US-09-825-751A-19  
Perfect score: 1425  
Sequence: 1 ctcctgggagacgcagccca.....tttccgcgcgttgccgcg 1425

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA:\*

- 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	987	69.3	1171	1	US-08-362-670B-25
2	987	69.3	1171	3	US-08-333-576C-25
3	987	69.3	1171	4	US-08-808-324-25
4	987	69.3	1171	5	PCT-US94-14030A-25
5	581.8	40.8	1046	1	US-08-362-670B-31
6	581.8	40.8	1046	3	US-08-333-576C-31
7	581.8	40.8	1046	4	US-08-808-324-31
8	581.8	40.8	1046	5	PCT-US94-14030A-31
9	412.2	28.9	530	1	US-08-581-529B-5
10	412.2	28.9	530	3	US-09-097-616-5
11	412.2	28.9	530	5	PCT-US94-07762-5
12	259.2	18.2	1345	1	US-08-362-670B-33
13	259.2	18.2	1345	3	US-08-333-576C-33
14	259.2	18.2	1345	4	US-08-808-324-33
15	259.2	18.2	1345	5	PCT-US94-14030A-33
16	253	17.8	2703	2	US-08-288-508C-1
17	253	17.8	2703	4	US-09-180-700-4
18	251.6	17.7	1207	1	US-08-362-670B-3
19	251.6	17.7	1207	3	US-08-333-576C-3
20	251.6	17.7	1207	5	US-08-289-222E-1
21	251.6	17.7	1207	4	US-09-054-526B-1
22	251.6	17.7	1207	4	US-08-808-324-3
23	251.6	17.7	1207	5	PCT-US94-14030A-3
24	249.2	17.5	519	2	US-08-581-528A-5
25	249.2	17.5	519	5	PCT-US94-07799-5
26	245.8	17.2	357	4	US-09-180-700-1
27	245	17.2	1203	1	US-08-362-670B-29

28	245	17.2	1203	3	US-08-333-576C-29	Sequence 29, Appl
29	245	17.2	1203	4	US-08-808-324-29	Sequence 29, Appl
30	245	17.2	1203	5	PCT-US94-14030A-29	Sequence 29, Appl
31	244.8	17.2	926	1	US-08-362-670B-1	Sequence 1, Appl
32	244.8	17.2	926	3	US-08-333-576C-1	Sequence 1, Appl
33	244.8	17.2	926	4	US-08-808-324-1	Sequence 1, Appl
34	244.8	17.2	926	5	PCT-US94-14030A-1	Sequence 1, Appl
35	239	16.8	2329	1	US-08-455-559-9	Sequence 9, Appl
36	239	16.8	2329	4	US-09-145-060-9	Sequence 9, Appl
37	239	16.8	2329	5	PCT-US94-00657-9	Sequence 9, Appl
38	232	16.3	272	1	US-08-362-670B-12	Sequence 12, Appl
39	232	16.3	272	3	US-08-333-576C-12	Sequence 12, Appl
40	232	16.3	272	4	US-08-808-324-12	Sequence 12, Appl
41	232	16.3	272	5	PCT-US94-14030A-12	Sequence 12, Appl
42	230	16.1	1233	1	US-08-362-670B-27	Sequence 27, Appl
43	230	16.1	1233	3	US-08-333-576C-27	Sequence 27, Appl
44	230	16.1	1233	4	US-08-808-324-27	Sequence 27, Appl
45	230	16.1	1233	5	PCT-US94-14030A-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1  
US-08-362-670B-25  
; Sequence 25, Application US/08362670B  
; Patent No. 5658882  
; GENERAL INFORMATION:  
; APPLICANT: Celeste, Anthony J.  
; APPLICANT: Wozney, John  
; APPLICANT: Rosen, Vicki A.  
; APPLICANT: Wolfman, Neil  
; APPLICANT: Thomsen, Gerald H.  
; APPLICANT: Melton, Douglas A.  
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENETICS INSTITUTE, INC.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/362,670B  
; FILING DATE: December 22, 1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lazar, Steven R.  
; REGISTRATION NUMBER: 32,618  
; REFERENCE/DOCKET NUMBER: 5202-D  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617 498-8260  
; TELEFAX: 617 876-5851  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1171 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; IMMEDIATE SOURCE:  
; CLONE: Human VL-1 protein  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 2..964  
; FEATURE:  
; NAME/KEY: mat\_peptide

LOCATION: 605..964  
US-08-362-670B-25

Query Match 69.3%; Score 987; DB 1; Length 1171;  
Best Local Similarity 100.0%; Pred. No. 2.7e-185;  
Matches 987; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 439 GATCTCTCGACACTCCTCTCCGGAGACAGAAGTATTTGTTGATGTGCCATGCTCTCA 498  
|||||  
Db 8 GATCTCTCGACACTCCTCTCCGGAGACAGAAGTATTTGTTGATGTGCCATGCTCTCA 67

QY 499 GACAAAGAAGAGCTGTGGCGCGGAGCTGCGGCTCTTTCCGACAGCGCCCTCAGCGCCC 558  
|||||  
Db 68 GACAAAGAAGAGCTGTGGCGCGGAGCTGCGGCTCTTTCCGACAGCGCCCTCAGCGCCC 127

QY 559 TGGGGGCCACACCGCGCGGCTCCAGCTGACGCTCTTCCCTTGGCTTTCGCCCTTACTG 618  
|||||  
Db 128 TGGGGGCCACACCGCGCGGCTCCAGCTGACGCTCTTCCCTTGGCTTTCGCCCTTACTG 187

QY 619 CTGGACGCGCGGACCTTGACCGCGAGGGGCGCGCGCGGCTGGGAAGTCTTCGAC 678  
|||||  
Db 188 CTGGACGCGCGGACCTTGACCGCGAGGGGCGCGCGCGGCTGGGAAGTCTTCGAC 247

QY 679 GTGTGGCAGGGCTGCGCCACCAAGCCCTGGAAGACGCTGTGCTTGGAGCTGCGGGCCGA 738  
|||||  
Db 248 GTGTGGCAGGGCTGCGCCACCAAGCCCTGGAAGACGCTGTGCTTGGAGCTGCGGGCCGA 307

QY 739 TGGGGCGAGCTGGACCGCGGGAGCGGAGCGCGCGCGGAGCCCGAGCAACCGCGG 798  
|||||  
Db 308 TGGGGCGAGCTGGACCGCGGGAGCGGAGCGCGCGGAGCCCGAGCAACCGCGG 367

QY 799 CCCCCGAGCTGCGGAGCTTGGGCTTCGCGGAGGCTCGGCTCCCGAGGAGCGGCG 858  
|||||  
Db 368 CCCCCGAGCTGCGGAGCTTGGGCTTCGCGGAGGCTCGGCTCCCGAGGAGCGGCG 427

QY 859 CTGCTGGTGGTATTACAGAGATCCCAAGCCCAAGAACCTGTTCCAGAGATGCGCGAGCAG 918  
|||||  
Db 428 CTGCTGGTGGTATTACAGAGATCCCAAGCCCAAGAACCTGTTCCAGAGATGCGCGAGCAG 487

QY 919 CTGGGCTCGCGGAGCTCGGCGCGCGCGCGGCGGAGGCTGCTGGCGCGCGCG 978  
|||||  
Db 488 CTGGGCTCGCGGAGCTCGGCGCGCGCGCGGCGGAGGCTGCTGGCGCGCGCG 547

QY 979 TGGGGCGCGCGGATGCCAGGCTTGGCTTCGCTTCGCGGCGCGCGCGCGCGCAG 1038  
|||||  
Db 548 TGGGGCGCGCGGATGCCAGGCTTGGCTTCGCTTCGCGGCGCGCGCGCGCGCAG 607

QY 1039 GCTTCGCCAGTCGCCATGGCAAGCGGCGACGCAAGAGTCCAGGCTACGCTGCAGCAAG 1098  
|||||  
Db 608 GCTTCGCCAGTCGCCATGGCAAGCGGCGACGCAAGAGTCCAGGCTACGCTGCAGCAAG 667

QY 1099 AAGCCCTCGCAGTGAACCTCAAGGAGCTGGGCTGGGAGGCTGATTCGCGCCCTG 1158  
|||||  
Db 668 AAGCCCTCGCAGTGAACCTCAAGGAGCTGGGCTGGGAGGCTGATTCGCGCCCTG 727

QY 1159 GAGTACGAGGCTATCAGTCGAGGCTGTATGCGACTTCCGCTCGCAGCTGGAG 1218  
|||||  
Db 728 GAGTACGAGGCTATCAGTCGAGGCTGTATGCGACTTCCGCTCGCAGCTGGAG 787

QY 1219 CCCACCAACCGCCATCATCCAGAGCTGTATGAACCTCCATGCCAGCCCGGCTCCACCCG 1278  
|||||  
Db 788 CCCACCAACCGCCATCATCCAGAGCTGTATGAACCTCCATGCCAGCCCGGCTCCACCCG 847

QY 1279 CCCAGTGTGGTGGCCCAACCAATGACTCCCATCAGATTCATACATCAGCGGGC 1338  
|||||  
Db 848 CCCAGTGTGGTGGCCCAACCAATGACTCCCATCAGATTCATACATCAGCGGGC 907

QY 1339 AATAATGTGGTCTAAGCAGTACGAGGACATGGTGGTGGAGTCTGCGGCTGCAGGTAG 1398  
|||||  
Db 908 AATAATGTGGTCTAAGCAGTACGAGGACATGGTGGTGGAGTCTGCGGCTGCAGGTAG 967

QY 1399 CGGTGCTTTCCGCGGCTTGGCCG 1425  
|||||

Db 968 CGGTGCTTTCCGCGCGCTTGGCCG 994

RESULT 2  
US-08-333-576C-25  
; Sequence 25, Application US/08333576C  
; Patent No. 6027919  
; GENERAL INFORMATION:  
; APPLICANT: Celeste, Anthony J.  
; APPLICANT: Wozney, John  
; APPLICANT: Rosen, Vicki A.  
; APPLICANT: Wolfman, Neil  
; APPLICANT: Thomsen, Gerald H.  
; APPLICANT: Melton, Douglas A.  
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: GENETICS INSTITUTE, INC.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/333,576C  
; FILING DATE: No. 6027919ember 2, 1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lazar, Steven R.  
; REGISTRATION NUMBER: 32,618  
; REFERENCE/DOCKET NUMBER: 5202-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617 498-8260  
; TELEFAX: 617 876-5851  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1171 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; IMMEDIATE SOURCE:  
; CLONE: Human VL-1 protein  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 2..964  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: 605..964  
US-08-333-576C-25

Query Match 69.3%; Score 987; DB 3; Length 1171;  
Best Local Similarity 100.0%; Pred. No. 2.7e-185;  
Matches 987; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 439 GATCTCTCGACACTCCTCTCCGGAGACAGAAGTATTTGTTGATGTGCCATGCTCTCA 498  
|||||  
Db 8 GATCTCTCGACACTCCTCTCCGGAGACAGAAGTATTTGTTGATGTGCCATGCTCTCA 67

QY 499 GACAAAGAAGAGCTGTGGCGCGGAGCTGCGGCTCTTTCCGACAGCGCCCTCAGCGCCC 558  
|||||  
Db 68 GACAAAGAAGAGCTGTGGCGCGGAGCTGCGGCTCTTTCCGACAGCGCCCTCAGCGCCC 127

QY 559 TGGGGGCCACACCGCGCGGCTCCAGCTGACGCTCTTCCCTTGGCTTTCGCCCTTACTG 618  
|||||  
Db 128 TGGGGGCCACACCGCGCGGCTCCAGCTGACGCTCTTCCCTTGGCTTTCGCCCTTACTG 187

QY 619 CTGGACGCGCGGACCTTGACCGCGAGGGGCGCGCGCGGCTGGGAAGTCTTCGAC 678



Db 188 CTGACCGCGGACCTTGGACCCGAGGGGCGCGCGCGCTGGGAAGTCTTCGAC 247  
QY 679 GTGTGGCAGGCGCTTGGCCACCAAGCCTTGAAGCAGTGTGCTTGGAGCTGCGGGCGCA 738  
Db 248 GTGTGGCAGGCGCTTGGCCACCAAGCCTTGAAGCAGTGTGCTTGGAGCTGCGGGCGCA 307  
QY 739 TGGGGGAGCTGTGACGCGCGGGGAGCGGAGCGCGCGCGGGGAGCCCAAGCAACGCCG 798  
Db 308 TGGGGGAGCTGTGACGCGCGGGGAGCGGAGCGCGCGCGGGGAGCCCAAGCAACGCCG 367  
QY 799 CCCCAGGAGCTGCGAGCTGTGGCTTCGCGCGAGGGTTCGGGCTCCCAAGAGCGGGCC 858  
Db 368 CCCCAGGAGCTGCGAGCTGTGGCTTCGCGCGAGGGTTCGGGCTCCCAAGAGCGGGCC 427  
QY 859 CTGCTGTGTATTACACAGATCCCAAGCAAGAACTGTTCAGAGATGCGCGAGCAG 918  
Db 428 CTGCTGTGTATTACACAGATCCCAAGCAAGAACTGTTCAGAGATGCGCGAGCAG 487  
QY 919 CTGGGCTGCGCGAGGCTGCGGGCGCGGGCGCGGGCGCGGGGCTGTGGCGCGCGCG 978  
Db 488 CTGGGCTGCGCGAGGCTGCGGGCGCGGGCGCGGGCGCGGGGCTGTGGCGCGCGCG 547  
QY 979 TCGGGCGCGCGAGTCCAGGCTTGGCTCCCTCGCGCGCGCGCGCGGGCGGGCGCGAG 1038  
Db 548 TCGGGCGCGCGAGTCCAGGCTTGGCTCCCTCGCGCGCGCGCGGGCGGGCGCGCGAG 607  
QY 1039 GCCTTGGCAGTCCCATGCAAGCGGCGCGGCAAGAGTCCAGGCTACGCTCAGCAAG 1098  
Db 608 GCCTTGGCAGTCCCATGCAAGCGGCGCGGCAAGAGTCCAGGCTACGCTCAGCAAG 667  
QY 1099 AAGCCCTGACGTAACTTCAAGAGCTGGGCTGGGAGCAGTGGATTATCGCGCGCGCTG 1158  
Db 668 AAGCCCTGACGTAACTTCAAGAGCTGGGCTGGGAGCAGTGGATTATCGCGCGCGCTG 727  
QY 1159 GAGTACGAGGCTTACGTCGAGGCTGTATCGAGCTTCCGCTGCGCTGCGACCTGGAG 1218  
Db 728 GAGTACGAGGCTTACGTCGAGGCTGTATCGAGCTTCCGCTGCGCTGCGACCTGGAG 787  
QY 1219 CCACCAACACGCCATCATCCAGAGCTGATGAATCCATGACCCCGGCTCCACCCCG 1278  
Db 788 CCACCAACACGCCATCATCCAGAGCTGATGAATCCATGACCCCGGCTCCACCCCG 847  
QY 1279 CCAGCTGTGCTGCGCCCAAAATGACTTCCCATCAGCAATTCATACATCGAGCGGGC 1338  
Db 848 CCAGCTGTGCTGCGCCCAAAATGACTTCCCATCAGCAATTCATACATCGAGCGGGC 907  
QY 1339 AATAATGTGTCTACAGCAGTACGAGGATGTTGGTGGAGTCTGCGGCTCGAGGTAG 1398  
Db 908 AATAATGTGTCTACAGCAGTACGAGGATGTTGGTGGAGTCTGCGGCTCGAGGTAG 967  
QY 1399 CGGTGCTTTCCCGCGCGCTTGGCCCG 1425  
Db 968 CGGTGCTTTCCCGCGCGCTTGGCCCG 994

RESULT 3

US-08-808-324-25  
; Sequence 25, Application US/08808324  
; Patent No. 6284872  
; GENERAL INFORMATION:  
; APPLICANT: Celeste, Anthony J.  
; APPLICANT: Wozney, John  
; APPLICANT: Rosen, Vicki A.  
; APPLICANT: Wolfman, Neil  
; APPLICANT: Thomsen, Gerald H.  
; APPLICANT: Melton, Douglas A.  
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENETICS INSTITUTE, INC.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge

; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/808,324  
; FILING DATE: Herewith  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lazar, Steven R.  
; REGISTRATION NUMBER: 32,618  
; REFERENCE/DOCKET NUMBER: 5202-D  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617 498-8260  
; TELEFAX: 617 876-5851  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1171 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; IMMEDIATE SOURCE:  
; CLONE: Human VL-1 protein  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 2..964  
; NAME/KEY: mat\_peptide  
; LOCATION: 605..964  
; US-08-808-324-25

Query Match 69.3%; Score 987; DB 4; Length 1171;  
Best Local Similarity 100.0%; Pred. No. 2.7e-185;  
Matches 987; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 439 GATCTCTCGACACTCTCTCCGGAGACAGAAATTTGTTGATGTGTCATGCTCTCA 498  
Db 8 GATCTCTCGACACTCTCTCCGGAGACAGAAATTTGTTGATGTGTCATGCTCTCA 67  
QY 499 GACAAAGAGAGCTGTGGCGCGGAGCTCGGCTCTTTCCGAGGCGCTCAGCGCCC 558  
Db 68 GACAAAGAGAGCTGTGGCGCGGAGCTCGGCTCTTTCCGAGGCGCTCAGCGCCC 127  
QY 559 TGGGGGCCACACAGCGGGCGGCTCCAGTGCAGCTCTTCCCTTTGCCCTTACTG 618  
Db 128 TGGGGGCCACACAGCGGGCGGCTCCAGTGCAGCTCTTCCCTTTGCCCTTACTG 187  
QY 619 CTGGACGCGGAGCCCTTGGACCCGCGAGGGGCGCGCGCGGCTGGGAAGTCTTCGAC 678  
Db 188 CTGGACGCGGAGCCCTTGGACCCGCGAGGGGCGCGCGCGGCTGGGAAGTCTTCGAC 247  
QY 679 GTGTGCAGGGCTTGGCCACCAAGCCTTGAAGCAGTGTGCTTGGAGCTGCGGGCGCA 738  
Db 248 GTGTGCAGGGCTTGGCCACCAAGCCTTGAAGCAGTGTGCTTGGAGCTGCGGGCGCA 307  
QY 739 TGGGGGAGCTGTGACGCGGGGAGCGGAGCGCGCGCGGGGAGCCCAAGCAACGCCG 798  
Db 308 TGGGGGAGCTGTGACGCGGGGAGCGGAGCGCGCGCGGGGAGCCCAAGCAACGCCG 367  
QY 799 CCCCAGGAGCTGCGAGCTTGGGCTTCGGCGCGAGGGTTCGGGCTCCCAAGAGCGGGCC 858  
Db 368 CCCCAGGAGCTGCGAGCTTGGGCTTCGGCGCGAGGGTTCGGGCTCCCAAGAGCGGGCC 427  
QY 859 CTGCTGTGTATTACACAGATCCCAAGCAAGAACTGTTCAGAGATGCGCGAGCAG 918  
Db 428 CTGCTGTGTATTACACAGATCCCAAGCAAGAACTGTTCAGAGATGCGCGAGCAG 487  
QY 919 CTGGGCTGCGCGAGGCTGCGGGCGCGGGCGCGGGGCTGTGGCGCGCGCGCG 978







QY	553	CGCGCGCGCTGGGAAGTC-TTCAGCTGTGGCAGGCGCTGCGCCACCAAGCCCTGGAAAG	711
Db	59	CGACCGCGCTGGGAAGTCTTTCAGCTGTGGCAGGCGCTGCGCCCTGACCTTGGAAAG	118
QY	712	CAGCTGTGCTTGAGCTGCGGCGCCGTCATGGGCGAGCTGAGCGCCGGGAGCGCAGGCG	771
Db	119	CAGCTGTGCTTGAGTGTGCGGCGACCTGTGGGTGAGCTGGACRCGGGGATACGGGGGCG	178
QY	772	CGGCGCGGGGACCCACGAAACCGCGCCCGCCGACCTGCGGAGTCTGGGCTTCGCGCGG	831
Db	179	CGCGCGAGGGTCCCCAGCAGCACCGCCTCTGGACCTGCGGAGTCTGGGCTTCGGTCGG	238
QY	832	AGGGTGGGCTCCCCAGAGCGGGCCTGCTGTGTGATTTCACAGATCCACGCGCAAG	891
Db	239	AGGGTGAGACCGCCAGAGGCGCGCCTGCTTGTGTGTTCACAGATCGACGCGCAAG	298
QY	892	AACCTGTTCCAGAGATGCGCAGCAGCTGGGCTCGGCGCAGGCTGCGGGCCCGGGCGCG	951
Db	299	AACCTGTTCACTGAGATGATGAGCAGCTGGGCTCTGCAGAGGCT-----GCG	346
QY	952	GGCGCGGAGGGTCTGTGGCGCGCGCTGCGGCGCCCGGATGCGAGCCTTGGCTGCC	1011
Db	347	GGAGCGGAGGGTCATGCCAGCGCGTGGGCTCCCCAGACCGGGTCTTGGCTGCC	406
QY	1012	TCGCGCGGCGCGGGGGGGCGACAGGCTTTCGCCAGTGCGCATATGGCAAGCGCACGGC	1071
Db	407	TCGCGCGGCGCGGGGGGAGCGCACCGCTTTCGCCAGCGCTTCACGGCAAGCGACATGGC	466
QY	1072	AAGAAGTCAGGCTACGCTGCAGCAAGAACCCCTGCAGCTGAACCTCAGGAGCTGGC	1131
Db	467	AAGAAGTCAGGCTGCGCTGCAGCAAGAACCCCTGCAGCTGAATTTAAGGAGTTAGGC	526
QY	1132	TGGGAGGACTTGGATTATCGCGCCCTTGGAGTACGAGGCTATCACTGCGAGGGTGTATGC	1191
Db	527	TGGGAGGACTTGGATTATCGCGCCCTAGAGTACGAGGCTATCACTGCGAGGGCTGTGC	586
QY	1192	GACTTCCGCTGGCTTCGACCTTGGAGGCCCAACACAGCGCATATCCAGAGCTGATG	1251
Db	587	GACTTTCGCTGGCTTCGACACCTTGAGGCCCAATAACCATGCATCATTCAGACGCTGATG	646
QY	1252	AACCTCATGGACCCGGGCTCACCCGCCCGCTGCTGCGTGGCCCAACCAATTCACCTCC	1311
Db	647	AACCTCATGGACCCGGGCTTCACCCCGCTAGCTGCTGCGTTCGCCCAACCAACTGACCTCC	706
QY	1312	ATCAGACTTCTATACATCGACCGGGCAATATGTGTCTACAAGCAGTACGAGGACATG	1371
Db	707	ATTAGCATCTGTACATCGACCGGGCAATATGTGTCTACAAGCAGTATGAGACATG	766
QY	1372	GTGGTGGAGTCTGGGCTGCAAGTAGCGTGCCCTTCCCGCGCGCTTGGCCCG	1425
Db	767	GTGGTGGAGTCTGGGCTGTAGTAGGCGGTG-CTGTCCCGCCACCTTGGCCAG	819

## RESULT 8

```

PCT-US94-14030A-31
; Sequence 31, Application PC/TUS9414030A
; GENERAL INFORMATION:
;
; APPLICANT: GENETICS INSTITUTE, INC.
; APPLICANT: PRESIDENT AND FELLOWS OF HARVARD COLLEGE
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```



```

; LENGTH: 530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: GDF-6
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 126..527
US-09-097-616-5

Query Match      28.9%; Score 412.2; DB 3; Length 530;
Best Local Similarity 87.8%; Pred. No. 1.2e-72;
Matches 477; Conservative 0; Mismatches 53; Indels 13; Gaps 2;

QY 856 GCCTCTGCTGGTGTATTACACAGATCCAGCGCAAGAACTGTTCCAGAGATGCGCGAG 915
Db 1 GCCTCTGCTGGTGTATTACACAGATCCAGCGCAAGAACTGTTCCAGATGCGATGAG 60

QY 916 CAGCTGGGCTCGGCCGAGGCTCGGGCGCGCGGCGCGGAGGGTCTGTGGCGCGCG 975
Db 61 CAGCTGGGCTCTGCAGAGGCT-----GCGGGAGCGGAGGGTCTATGCGCAGCG 108

QY 976 CCGTCGGGCGCGCGGATGCCAGGCTTGGCTGCCCTCGCGCGCGCGCGCGCGCGCG 1035
Db 109 CCGTCGGGCG-TCCAGAGCGCGGGTCTTGGCTGCCCTCGCGCGCGCGCGCGCGCGCG 167

QY 1036 ACGGCGCTTCGCCAGTCCCATGGCAAGCGCGCACGCAAGATCCAGGCTACGCTGCAGC 1095
Db 168 ACGGCGCTTCGCCAGCGCTCAGCGCAAGCGCACATGGCAAGATCCAGGCTGCGCTGCAGC 227

QY 1096 AAGAACCCCTCGCACGTGAACCTCAAGAGCTGGGCTGGGAGAGCTGGATTATCGCGGCC 1155
Db 228 AAGAACCCCTCTGCACGTGAATTTAAGGAGTAAAGCTGGGAGCTGGATTATCGCGGCC 287

QY 1156 CTGGAGTACGAGGCTTACTCTCGAGGGTGTATGCGACTTCCCGTGGCTGCGCTCGCACCTG 1215
Db 288 CTAGAGTACGAGGCTTACTCTCGAGGGCGTGTGCGACTTTCCTGGCTGCGCTCGCACCTT 347

QY 1216 GAGCCCAACCAACCGCCATCATCCAGAGCGCTGATGAAGTCCATGACCCCGGCTCCACC 1275
Db 348 GAGCCCACTAACCATGCCATCATTCAGAGCGCTGATGAAGTCCATGACCCCGGCTCCACC 407

QY 1276 CGGCCAGCTGTGCTGCCGCCACCAATTTGACTCCCATCAGCAATTTCTATACATCGACGG 1335
Db 408 CGGCCAGCTGTGCTGCCGCCACCAATTTGACTCCCATCAGCAATTTCTATACATCGACGG 467

QY 1336 GGCATAATGTGGTCTACAGAGAGTACGAGGACATGGTGGTGGAGTCTGCGGCTCGAGG 1395
Db 468 GGCATAATGTAGTCTACAGAGAGTATGAGGACATGGTGGTGGAGTCTGCGGCTGTAGG 527

QY 1396 TAG 1398
Db 528 TAG 530
```

```

RESULT 11
PCT-US94-07762-5
; Sequence 5; Application PC/TUS9407762
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-6
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07762
; FILING DATE: 08-JUL-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: TUMARKIN, LISA A., PH.D.
; REGISTRATION NUMBER: P-38,347
; REFERENCE/DOCKET NUMBER: FD2349
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: GDF-6
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 126..527
PCT-US94-07762-5

Query Match      28.9%; Score 412.2; DB 5; Length 530;
Best Local Similarity 87.8%; Pred. No. 1.2e-72;
Matches 477; Conservative 0; Mismatches 53; Indels 13; Gaps 2;

QY 856 GCCTCTGCTGGTGTATTACACAGATCCAGCGCAAGAACTGTTCCAGAGATGCGCGAG 915
Db 1 GCCTCTGCTGGTGTATTACACAGATCCAGCGCAAGAACTGTTCTACTGAGATGCGATGAG 60

QY 916 CAGCTGGGCTCGGCCGAGGCTCGGGCGCGCGGCGCGGAGGGTCTGTGGCGCGCG 975
Db 61 CAGCTGGGCTCTGCAGAGGCT-----GCGGGAGCGGAGGGTCTATGCGCAGCG 108

QY 976 CCGTCGGGCGCGCGGATGCCAGGCTTGGCTGCCCTCGCGCGCGCGCGCGCGCGCGCG 1035
Db 109 CCGTCGGGCG-TCCAGAGCGCGGGTCTTGGCTGCCCTCGCGCGCGCGCGCGCGCGCGCG 167

QY 1036 ACGGCGCTTCGCCAGTCCCATGGCAAGCGCGCACGCAAGATCCAGGCTACGCTGCAGC 1095
Db 168 ACGGCGCTTCGCCAGCGCTCAGCGCAAGCGCACATGGCAAGATCCAGGCTGCGCTGCAGC 227

QY 1096 AAGAACCCCTCGCACGTGAACCTCAAGAGCTGGGCTGGGAGAGCTGGATTATCGCGGCC 1155
Db 228 AAGAACCCCTCTGCACGTGAATTTAAGGAGTAAAGCTGGGAGCTGGATTATCGCGGCC 287

QY 1156 CTGGAGTACGAGGCTTACTCTCGAGGGTGTATGCGACTTCCCGTGGCTGCGCTCGCACCTG 1215
Db 288 CTAGAGTACGAGGCTTACTCTCGAGGGCGTGTGCGACTTTCCTGGCTGCGCTCGCACCTT 347

QY 1216 GAGCCCAACCAACCGCCATCATCCAGAGCGCTGATGAAGTCCATGACCCCGGCTCCACC 1275
Db 348 GAGCCCACTAACCATGCCATCATTCAGAGCGCTGATGAAGTCCATGACCCCGGCTCCACC 407

QY 1276 CGGCCAGCTGTGCTGCCGCCACCAATTTGACTCCCATCAGCAATTTCTATACATCGACGG 1335
Db 408 CGGCCAGCTGTGCTGCCGCCACCAATTTGACTCCCATCAGCAATTTCTATACATCGACGG 467

QY 1336 GGCATAATGTGGTCTACAGAGAGTACGAGGACATGGTGGTGGAGTCTGCGGCTCGAGG 1395
Db 468 GGCATAATGTAGTCTACAGAGAGTATGAGGACATGGTGGTGGAGTCTGCGGCTGTAGG 527

QY 1396 TAG 1398
Db 528 TAG 530
```







```

; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: HUMAN V1-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 138..1301
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 990..1301
US-08-808-324-33

Query Match      18.2%; Score 259.2; DB 4; Length 1345;
Best Local Similarity 57.8%; Pred. No. 1.4e-42;
Matches 550; Conservative 0; Mismatches 378; Indels 24; Gaps 4;

QY 464 GACAGAAATATTGTTGATGTCCTCATGCTCTCAGACAAAGAGCTGGTGGCGCGG 523
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 364 GCACAGATTCCTGTTGCGAGTGTCCAGCTTAACAGCAGACAGAGGTGGTGGTGC 423
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 524 AGCTGCGGCTCTTTCCAGCGCCCTCAGCGCCCTGGGGCCACACAGCGGGCGCTCC 583
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 424 AGCTGCGGCTGTCGCGCGGGATCTCCAGATC--GGCCCCAGGAGCTGACTTCTCC 481
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 584 AGCTGAGCTCTCCCTTTCGCTTTCGCCCTACTGCTGGACGCGGACCCCTGGACCCGC 643
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 482 -----GCCGTTGCTGCTGCTGCACGTGCCCGGGCCCGCGAGCGCCACGCTGC 534
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 644 AGGGGCGCGCGCGGCTGGAGTCTTCGACGTGTGGCAGGCGCTGCCACCAGC 703
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 535 TGACTCGCGGCGAGTGAGCCCTAGTCGCTGAGCGTGGAGGCGTTCGACGTGGCGG 594
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 704 CTGGAAGCAGCTGTGCTTGGAGCTCGCGCGCGCATGGGCGGAGCTGGACGCGGGGAG 763
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 595 ACGCCATGAGCGCACCGCTGTAACCGCGCCCCCGCGCGTTCGCTTCTGCTGC 654
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 764 CCGAGCGCGCGGGGAGACCCAGAACCGCGCGCCCGCCCGGAGCTCGGGT 823
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 655 GGCAGTGTGCAGCGCGGCTGCGGAGCCCTTGGCAGCTGCGGCGACTGGGCTTCGGCTGC 714
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 824 TCGCGGAGGCTGGGCGCTCCCGAGGAGCGGCGCTGCTGTGTTATTCACAGATCC 883
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 715 CGGGCGGAGGGGCTGTCGGGAGAGGAGCGCGCGGTGCTAGTCTCTCCCGCAGC 774
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 884 AGCGCA---AGAACCTGTTTCGAGAGATCGCGAGCAGCTGGGCTCGCGCGAGGCTGC-- 938
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 775 AGAGGAAGAGAGCTTATTCCGSGAGATCCGCGCCAGCGCGCGCTCGGGCGCGCTC 834
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 939 -----GGGCGCGCGCGCGCGGAGGGTGTGTGGCGCGCGCGCTCGGGCGCGCC 988
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 835 TGGCCTCAGAGCGGCTGCCCGACCCAGGAACCGGACCGCTCGGCAAGGGCAGTCA 894
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 989 CGGATGCCAGGCTTGGCTGCGCTCCCGCGCGCGCGCGCGCGCGCGCTTCGCCA 1048
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 895 GCGGCGGAGGAGGAGGAGCGGCTTGGCGGAGCGGAGCGGAGCGGAGCGGAGCGCGG 954
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 1049 GTCCCATGGCAAGCGGACGCGCAAGATCGCTGACGAGCAAGAACGCCCTTCG 1108
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 955 GGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1014
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 1109 AGCTGAATTCAGAGCTGGGCTGGGAGGAGTATTCGCGCCCTTCGAGTACGAGG 1168
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 1015 AGCTGAGCTTCAAGAGCTCGGCTGGGAGCTGATGATCATCGCGCGCTGACTACGAGG 1074
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 1169 CTTATCACTGCGAGGCTGTATCGACTTCCGCTCGGCTCGGCTCGGAGCCCAACC 1228
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 1075 CGTACACTGCGAGGCGCTTTCGACTTCCCTTTCGCTGCGGAGCTCGAGCCCAACC 1134
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 1229 AGCCCATCATCAGAGCTGATGAATCATGAGACCCGGGTTCACCCCGCGCGAGCTGCT 1288
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 1135 ATGCCATCATTCAGAGCTGTCTCAACTCCATGGCAGCAGCGCGCGCGCTCTCTGCT 1194
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 1289 GGTGCCCAACCAATTGACTCCCATCAGATTCATATCATGACCGCGGCGCAATATGTTG 1348
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```

```

Db 1195 GTGTGCCAGCGCGCTCAGCCCCATCAGCATCTCTACATCAGCCGCCCAACAAGTTG 1254
QY 1349 TCTACAAGCAGTACGAGGACATGGTGGTGGAGTCTGTCGGGCTGCAGGTAGCG 1400
Db 1255 TCTACAAGCAATACGAGGACATGGTGGTGGAGGCTTGGGCTGCAGGTAGCG 1306

RESULT 15
PCT-US94-14030A-33
; Sequence 33, Application PC/TUS9414030A
; GENERAL INFORMATION:
; APPLICANT: GENETICS INSTITUTE, INC.
; APPLICANT: PRESIDENT AND FELLOWS OF HARVARD COLLEGE
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/14030A
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/164,103
; FILING DATE: 07-DEC-1993
; APPLICATION NUMBER: US 08/217,780
; FILING DATE: 25-MAR-1994
; APPLICATION NUMBER: US 08/333,576
; FILING DATE: 02-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202D-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1345 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: HUMAN V1-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 138..1301
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 990..1301
PCT-US94-14030A-33
```

```

Query Match      18.2%; Score 259.2; DB 5; Length 1345;
Best Local Similarity 57.8%; Pred. No. 1.4e-42;
Matches 550; Conservative 0; Mismatches 378; Indels 24; Gaps 4;

QY 464 GACAGAAATATTGTTGATGTCCTCATGCTCTCAGACAAAGAGCTGGTGGCGCGG 523
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 364 GCACAGATTCCTGTTGCGAGTGTCCAGCTTAACAGCAGACAGAGGTGGTGGTGC 423
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 524 AGCTGCGGCTCTTTTCGCGAGGCGGCTCAGGCGGCTTCGCGGCGCCACAGCGGGCGCTCC 583
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 01:02:45 ; Search time 86 Seconds  
(without alignments)  
6275.472 Million cell updates/sec

Title: US-09-825-751A-19

Perfect score: 1425

Sequence: 1 cctctggggagacgcagcca.....tttccgcgcgttgccgcg 1425

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 6711156

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_NA.\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	987	69.3	1171	9	US-09-945-182-25
2	928.2	65.1	1308	10	US-09-730-772-12
3	928.2	65.1	1308	10	US-09-735-849-12
4	581.8	40.8	1046	9	US-09-945-182-31
5	259.2	18.2	1345	9	US-09-945-182-33
6	253	17.8	2703	8	US-08-981-490B-2
7	251.6	17.7	1207	9	US-09-945-182-3
8	251.4	17.6	2341	10	US-09-730-772-11
9	251.4	17.6	2341	10	US-09-735-849-11
10	245.8	17.2	357	8	US-08-945-459A-4
11	245.8	17.2	357	8	US-09-068-253-1
12	245	17.2	1203	9	US-09-945-182-29
13	244.8	17.2	926	9	US-09-945-182-1
14	239	16.8	2329	9	US-09-880-708-9
15	237	16.6	475	10	US-09-833-381-582
16	232	16.3	272	9	US-09-945-182-12
17	230	16.1	1233	9	US-09-945-182-27
18	208.8	14.7	3585	9	US-09-945-182-9
19	195.2	13.7	272	9	US-09-945-182-10

20	172.8	12.1	272	9	US-09-945-182-14
21	144	10.1	1346	10	US-09-784-911-11
22	144	10.1	1352	10	US-09-784-911-13
23	138.4	9.7	557	10	US-09-765-527-248
24	138.4	9.7	557	10	US-09-760-397-1
25	138.4	9.7	1822	8	US-08-822-186-1
26	138.4	9.7	1822	8	US-08-937-755-1
27	138.4	9.7	1822	9	US-09-982-543A-9
28	138.4	9.7	1822	10	US-09-045-331-1
29	138.4	9.7	1822	10	US-09-828-607-1
30	138.4	9.7	1878	10	US-09-954-456-1824
31	136.2	9.6	610	10	US-09-760-397-7
32	136.2	9.6	613	10	US-09-765-527-256
33	136.2	9.6	613	10	US-09-765-527-260
34	136.2	9.6	646	10	US-09-760-397-13
35	136.2	9.6	658	10	US-09-765-527-254
36	136.2	9.6	661	10	US-09-765-527-262
37	131	9.2	1350	10	US-09-784-911-7
38	131	9.2	1362	10	US-09-784-911-3
39	130.8	9.2	1353	10	US-09-784-911-9
40	114.8	8.1	1926	12	US-10-044-090-28
41	114.4	8.0	1224	10	US-09-749-728B-70
42	111	7.8	1547	9	US-10-044-716-1
43	111	7.8	14759	10	US-09-952-360-1
44	107.2	7.5	1289	10	US-09-784-911-5
45	106.8	7.5	2003	10	US-09-764-869-1535

ALIGNMENTS

RESULT 1  
US-09-945-182-25  
Sequence 25, Application US/09945182  
Patent No. US20020160494A1  
GENERAL INFORMATION:  
APPLICANT: Celeste, Anthony J.  
Wozney, John  
Rosen, Vicki A.  
Wolfman, Neil  
Thomsen, Gerald H.  
Melton, Douglas A.  
TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENETICS INSTITUTE, INC.  
STREET: 87 CambridgePark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/945,182  
Filing DATE: 31-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/808,324  
Filing DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Lazar, Steven R.  
REGISTRATION NUMBER: 32,618  
REFERENCE/DOCKET NUMBER: 5202-D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617 498-8260  
TELEFAX: 617 876-5851  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1171 base pairs

TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
IMMEDIATE SOURCE:  
CLONE: Human VL-1 protein  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2..964  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 605..964  
SEQUENCE DESCRIPTION: SEQ ID NO: 25:  
US-09-945-182-25

Query Match 69.3%; Score 987; DB 9; Length 1171;  
Best Local Similarity 100.0%; Pred. No. 5.7e-192;  
Matches 987; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 439 GATCTCTCGCACACTCTCTCCGGAGACAGAGTATTTGTTGATGTGCTCCATGCTCTCA 498  
|||||  
Db 8 GATCTCTCGCACACTCTCTCCGGAGACAGAGTATTTGTTGATGTGCTCCATGCTCTCA 67  
|||||

QY 499 GACAAAGAGAGCTGTGGGCGGAGCTGCGGCTCTTTGCGCAGGCGCCCTCAGCGCCC 558  
|||||  
Db 68 GACAAAGAGAGCTGTGGGCGGAGCTGCGGCTCTTTGCGCAGGCGCCCTCAGCGCCC 127  
|||||

QY 559 TGGGGGCCACACCGCGGCTCCACGTGACGCTCTTCCCTTGCTTTCGCCCTACTG 618  
|||||  
Db 128 TGGGGGCCACACCGCGGCGCTCCACGTGACGCTCTTCCCTTGCTTTCGCCCTACTG 187  
|||||

QY 619 CTGGAGCGCGGACCTTGAACCGGAGGCGCGCGCGCGCTGGGAAGTCTTCGAC 678  
|||||  
Db 188 CTGGAGCGCGGACCTTGAACCGGAGGCGCGCGCGCGCTGGGAAGTCTTCGAC 247  
|||||

QY 679 GTGTGCGAGGCTGCGCCACACGCGCTGGAAGACGCTGTGCTTGAGCTGCGGGCGCA 738  
|||||  
Db 248 GTGTGCGAGGCTGCGCCACACGCGCTGGAAGACGCTGTGCTTGAGCTGCGGGCGCA 307  
|||||

QY 739 TGGGGCGAGCTGACCGCGGAGGCGGAGGCGCGCGCGGAGCCCGACGACCGCG 798  
|||||  
Db 308 TGGGGCGAGCTGACCGCGGAGGCGGAGGCGCGCGCGGAGCCCGACGACCGCG 367  
|||||

QY 799 CCCCCGAGCTGCGGAGCTTGGGCTTCGCGGAGGCTGCGGCTCCCGAGGAGCGGCG 858  
|||||  
Db 368 CCCCCGAGCTGCGGAGCTTGGGCTTCGCGGAGGCTGCGGCTCCCGAGGAGCGGCG 427  
|||||

QY 859 CTGCTGTGTATTCACAGATCCAGCGAGAACTGTTCGACAGATGCGCGAGCAG 918  
|||||  
Db 428 CTGCTGTGTATTCACAGATCCAGCGAGAACTGTTCGACAGATGCGCGAGCAG 487  
|||||

QY 919 CTGGGCTCGGCGAGGCTGCGGCGCGGCGGCGGCGGAGGCTGCGGCGCGCGCG 978  
|||||  
Db 488 CTGGGCTCGGCGAGGCTGCGGCGCGGCGGCGGCGGAGGCTGCGGCGCGCGCG 547  
|||||

QY 979 TCGGGCGCGCGGATGCCAGGCTTGGCTGCCCTCGCGGCGCGCGCGCGCGCGCAG 1038  
|||||  
Db 548 TCGGGCGCGCGGATGCCAGGCTTGGCTGCCCTCGCGGCGCGCGCGCGCGCGCAG 607  
|||||

QY 1039 GCTTTCGAGTGCATGCGAGCGGCGAGGCGAGAGTCCAGGCTACGCTGCAGCAAG 1098  
|||||  
Db 608 GCTTTCGAGTGCATGCGAGCGGCGAGGCGAGAGTCCAGGCTACGCTGCAGCAAG 667  
|||||

QY 1099 AAGCCCTCGACGTGACATCAAGAGCTGGGCTGGGAGACTGGATTATCGGCCCTTG 1158  
|||||  
Db 668 AAGCCCTCGACGTGACATCAAGAGCTGGGCTGGGAGACTGGATTATCGGCCCTTG 727  
|||||

QY 1159 GAGTACGAGGCTATCACTGCGAGGCTGTATGCGACTTCCGCTGCGCTCGCACCTGGAG 1218  
|||||  
Db 728 GAGTACGAGGCTATCACTGCGAGGCTGTATGCGACTTCCGCTGCGCTCGCACCTGGAG 787  
|||||

QY 1219 CCGACCAACCGCATATCCAGAGCTGTGAACTCCATGACGCCCGGCTCCACCCCG 1278  
|||||

Db 788 CCACCAACACGCCCATCCAGAGCTGATGAATCCATGACCCCGGCTCCACCCG 847  
QY 1279 CCAGCTGCTGCTGCCGCCCAAAATTTGACTCCCATCAGATTCTATACATCGAGCGGGC 1338  
|||||  
Db 848 CCAGCTGCTGCTGCCGCCCAAAATTTGACTCCCATCAGATTCTATACATCGAGCGGGC 907  
|||||

QY 1339 AATAATGTGTCTTACAAAGCAGTACGAGGACATGTTGTTGAGTCTGTCGGCTCAGGTAG 1398  
|||||  
Db 908 AATAATGTGTCTTACAAAGCAGTACGAGGACATGTTGTTGAGTCTGTCGGCTCAGGTAG 967  
|||||

QY 1399 CGGTGCTTTCCGCGCGCTTGGCCG 1425  
|||||  
Db 968 CGGTGCTTTCCGCGCGCTTGGCCG 994  
|||||

RESULT 2  
US-09-730-772-12  
; Sequence 12, Application US/09730772  
; Patent No. US20010011131A1  
; GENERAL INFORMATION:  
; APPLICANT: Layten, Frank P.  
; APPLICANT: Moos, Jr., Malcolm  
; APPLICANT: Chang, Steven Chao-Huan  
; TITLE OF INVENTION: CARTILAGE-DERIVED MORPHOGENETIC  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: Knobbe, Martens, Olson & Bear  
; STREET: 620 Newport Center Drive, 16th Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: U.S.A.  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/730,772  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/836,081  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bartfield, Neil S  
; REGISTRATION NUMBER: 39,901  
; REFERENCE/DOCKET NUMBER: NIH099.001APC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-235-8550  
; TELEFAX: 619-235-0176  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1308 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-09-730-772-12

Query Match 65.1%; Score 928.2; DB 10; Length 1308;  
Best Local Similarity 86.5%; Pred. No. 5e-180;  
Matches 1136; Conservative 0; Mismatches 123; Indels 54; Gaps 8;

QY 128 CGTCCGCCAGCTGGGTTCACCAAGGGCATCGAAGCGCCCAAGGAAGCAAGATCGAGC 187  
|||||  
Db 5 CGTCCGCCAGCTGGGTTCACCAAGGGCATCGAAGCGCCCAAGGAAGCAAGATCGAGC 64  
|||||

QY 188 GGGCGCGCGCGACAGTACGCGGCGCGGAG-----GGCCAGGAACACAGCGCGC 238  
|||||  
Db 65 GGGCGCGCGAGAGATGCGACGCGCGCGGAGCCCTGTGATGCCAGGAGCCCCCGCGA 124  
|||||

QY 239 GGCTTCAGGACGAACCC-----CGGGCTCAGCAGCCCCGGGGCGAGAGCCGCCAG 289  
Db 125 GGCGCAGGAGGAGCCCGGCGGCGCCACAGCAGCCTGAAGCTCGGAGCCTCCCG 184  
QY 290 GCAGGGTCCGGCGTGGTGGCCCGCAGAGTACATGTGTCAATCTACAGGACTTACTCCA 349  
Db 185 GCAGGGCCCGCGCTTGGTGGCCCGCAGAGTACATGTGTCAATCTACAGGACTTACTCCA 244  
QY 350 TCAGTGAAGAGCTGGGATCAATGTCAGCTTTTCCAGTCTTCCAGTCCGCTAATACGA 409  
Db 245 TCAGGAGAGAGCTGGGATCAATGTCAGCTTTTCCAGTCTTCCAGTCCGCTAATACGA 304  
QY 410 TCACAGCTTTGTAGACAGGGGACTAGACGATCTCTCGCACACTCTCTCCGGAGACAGA 469  
Db 305 TCAGTGAAGAGCTGGGATCAATGTCAGCTTTTCCAGTCTTCCAGTCCGCTAATACGA 364  
QY 470 AGTATTTGTTGATGTGTCATCTCTCAGACAAAGAGTGTGGGGCGGAGGTGC 529  
Db 365 AGTATTTGTTGATGTGTCATCTCTCAGACAAAGAGTGTGGGGCGGAGGTGC 424  
QY 530 GGCTCTTTCGCCAGGCGCTCAGCGCCCTGGGGGCCACACAGCGGGCGCTCACGTGC 599  
Db 425 GGCTCTTTCGCCAGGCGCTCAGCGCCCTGGGGGCCACACAGCGGGCGCTCACGTGC 479  
QY 590 AGCTCTTTCCTTCCCTTTCGCCCTTACTGCTGACGCGCGGACCCCTGGACCCGAGGGG 649  
Db 480 AGCTCTTTCCTTCCCTTTCGCCCTTACTGCTGACGCGCGGACCCCTGGACCCGCA-GGG 538  
QY 650 CGCGCGGGCGGCTGGGAAGTCTTGCAGTGTGGAGGCGCTTGGCGCACACCCCTTGA 709  
Db 539 CGCGCGGGCGGCTGGGAAGTCTTGCAGTGTGGGGGCGCTTGGCGCACACCCCTTGA 598  
QY 710 AGCAGCTGTGTTGAGCTCGCGGCGCATGG--GSCGAGCTGGAGCGCGGGGAGCGCG 766  
Db 599 AGCAGCTGTGTTGAGCTTCGGGCGCGCTGGGGCGGAGCGCGGGGAGCGCG 658  
QY 767 AGCGCGCGGCGGAGACCCAGCAACCGCGCGCGGACCTTGGGAGTCTGGGGTTCG 826  
Db 659 AGCGCGCGGCGGAGACCCAGCAACCGCGCGCGGACCTTGGGAGTCTGGGGTTCG 718  
QY 827 GCAGGAGTGGCGCTTCCAGAGGGGCGCTTGTGTGTGTTATTCACAGATCCAGC 886  
Db 719 GCAGGAGTGGCGCTTCCAGAGGGGCGCTTGTGTGTGTTATTCACAGATCCAGC 778  
QY 887 GCAAGAACCTGTTTCGAGAGATCGCGAGCAGCTGGGCTCG--CCGAGGCTCGGGCC 943  
Db 779 GCAAGAACCTGTTTCGAGAGATCGCGAGCAGCTGGGCTCGGGGAGCGGAGTGGTGGCC 838  
QY 944 CGGGCGGCGGCGCGAGGGTCTGT-----GCCCGCGCGCTCGG 982  
Db 839 CGGGTGTGGGGCGGAGGGTCTGGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 898  
QY 983 GCAGCGCGGATGCGAGGCTTGGCTTGGCTTGGCGCGCGCGCGCGCGCGCGCGCGCT 1042  
Db 899 GCAGCGCGGAGCTGGGCTTGGCTTGGCTTGGCGCGCGCGCGCGCGCGCGCGCT 955  
QY 1043 TCAGCAGTCCCATGTCAGCGCGCAGCGCAAGAGTCCAGGCTACGTCGACGCAAGAC 1102  
Db 956 TCAGCAGTCCCATGTCAGCGCGCAGCGCAAGAGTTCGAGGCTGCGCTGCGAGCAAGAC 1015  
QY 1103 CCCTGCAGCTGAATTCAGAGAGCTGGGCTGGGACGACTGGATTATCGCGCGCTGGAGT 1162  
Db 1016 CCCTGCAGCTGAATTCAGAGAGCTGGGCTGGGACGACTGGATTATCGCGCGCTGGAGT 1075  
QY 1163 ACAGGCGCTATCTAGGAGGTGTATGGAGCTTCCCGTTCGCGCTCGACCTGGAGCCCA 1222  
Db 1076 ACAGGCGCTATCTAGGAGGTGTATGGAGCTTCCCGTTCGCGCTCGACCTGGAGCCCA 1135  
QY 1223 CCAACACGCGCATCTACAGAGCTGATGAACCTTCCATGACCCCGGCTCCACCCCGCCCA 1282  
Db 1136 CCAACACGCGCATCTACAGAGCTGATGAACCTTCCATGACCCCGGCTCCACCCCGCCCA 1195

QY 1283 GGTGCTGCTGCCGCGCCACCAATTTGACTCCCATCAGCATTTCTATATCAGCGCGGCAATA 1342  
Db 1196 GGTGCTGCTGCCGCGCCACCAATTTGACTCCCATCAGCATTTCTATATCAGCGCGGCAATA 1255  
QY 1343 ATGTGCTCTACAGAGCTAGGAGACATGTGTGGAGTCTGCTGGGCTGCAGG 1395  
Db 1256 ATGTGCTCTACAGAGCTAGGAGACATGTGTGGAGTCTGCTGGGCTGCAGG 1308  
RESULT 3  
US-09-735-849-12  
; Sequence 12, Application US/09735849  
; Patent No. US20010037017A1  
; GENERAL INFORMATION:  
; APPLICANT: Luyten, Frank P.  
; APPLICANT: Moos, Jr., Malcolm  
; APPLICANT: Chang, Steven Chao-Huan  
; TITLE OF INVENTION: CARTILAGE-DERIVED MORPHOGENETIC  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 620 Newport Center Drive, 16th Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: U.S.A.  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/735,849  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/836,081  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bartfeld, Neil S  
; REGISTRATION NUMBER: 39,901  
; REFERENCE/DOCKET NUMBER: NIH099,001APC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-235-8550  
; TELEFAX: 619-235-0176  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1308 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna  
US-09-735-849-12  
Query Match 65.1%; Score 928.2; DB 10; Length 1308;  
Best Local Similarity 86.5%; Pred. No. 5e-180;  
Matches 1136; Conservative 0; Mismatches 123; Indels 54; Gaps 8;  
QY 128 CGTCCCGCGAGCTGGGTTCACCAAGGCGATGCGAGCCCGCAAGGAGCAAGATGCAGC 187  
Db 5 CGTCCCGCGAGCTGGGTTCGCGCCAGGGAATGCAACGCCCAAGGAAGGAAGATGCCGC 64  
QY 188 GGGCGCGCGCGACAGTGAAGCGCGGGCGGGAG-----GGCCAGGAACACAGCCGC 238  
Db 65 GGGCGCGCGAGAGAATGCCAGCGCGCGGGAGCCCTGGATCGCAGGAGCCCGCGCGA 124  
QY 239 GGCTTCAGGACGAACCC-----CGGGCTCAGCAGCCCCGGGGCGCAGAGCCCGCAG 289  
Db 125 GGCGCAGGAGGAGGAGCCCGAGCGCGCGCGCACAGCGCTGAAGCTCGGGAGCGCTCCG 184  
QY 290 GCAGGGTCCGCGGTGGTGGCGCGGATGCTGTCAATCTACAGGACTTACTCCA 349





Db 179 CGCGGAGGGTCCCGACAGCCACCGCCTCTGGACCTGGGAGCTGGGCTCGG 238  
QY 832 AGGTGCGGCTCCCGAGGCGGGCCCTGCTGTGTATTCACAGATCCCGCGCAAG 891  
Db 239 AGGTGAGAGCGCGCGAGGAGCGCCCTGCTTGTAGTTTACAGATCGCGCGCAAG 298  
QY 892 AACCTGTTCCAGAGATGCGCGAGAGCTGGGCTCGCGCGAGGCTGGGGCCCGGGCGG 951  
Db 299 AACCTGTTCACTAGATGATGATGATGATGATGATGATGATGATGATGATGAT 346  
QY 952 GCGCGCGAGGGTCTGTGGCGCGCGCTGGGCGCGCCCGAGTCCAGGCTTGGCTGCC 1011  
Db 347 GGAGCGAGGGGTCTATGTCCAGCGCGCTGGGCTCCCGACACCGGGTCTTGGCTGCC 406  
QY 1012 TCGCCCGCGCGCGGGCGGCGACGCGCTTCGCCAGTGGCCATGGCAAGCGCGCACGGC 1071  
Db 407 TCGCGCGCGCGCGGGCGGCGACCGCGCTTCGCCAGCGCTCACGCAAGCGACATGGC 466  
QY 1072 AAGAAAGTCCAGGCTAGCTGCACAGAGAGCCCTTCGACGTGAACCTTCAAGAGCTGGC 1131  
Db 467 AAGAAAGTCCAGGCTAGCTGCACAGAGAGCCCTTCGACGTGAACCTTCAAGAGCTGGC 526  
QY 1132 TGGAGAGCTGAGTATTCGCGCGCCCTGGAGTACGAGGCTTATCAGTGGAGGTGTATGC 1191  
Db 527 TGGAGAGCTGAGTATTCGCGCGCCCTGGAGTACGAGGCTTATCAGTGGAGGTGTATGC 586  
QY 1192 GACTTCCCGTGGCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1251  
Db 587 GACTTCCCGTGGCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 646  
QY 1252 AACTCATGAGACCGCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1311  
Db 647 AACTCATGAGACCGCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 706  
QY 1312 ATCAGACTTATATACATCGAGCGGCGCAATAATGTCTTACAGAGTACGAGGACATG 1371  
Db 707 ATTAGCATCTGTACATCGAGCGGCGCAATAATGTCTTACAGAGTATGAGGACATG 766  
QY 1372 GTGGTGGAGTCGCGGCTGCAGGTAGCGGTGCTTTCGCGCGCGCGCGCGCGCGCG 1425  
Db 767 GTGGTGGAGTCGCGGCTGCAGGTAGCGGTGCTTTCGCGCGCGCGCGCGCGCGCG 819

RESULT 5

US-09-945-182-33  
; Sequence 33, Application US/09945182  
; Patent No. US20020160494A1  
; GENERAL INFORMATION:  
; APPLICANT: Celeste, Anthony J.  
; Wozney, John  
; Rosen, Vicki A.  
; Wolfman, Neil  
; Thomsen, Gerald H.  
; Melton, Douglas A.  
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENETICS INSTITUTE, INC.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/945,182  
; FILING DATE: 31-Aug-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/808,324  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Lazar, Steven R.  
REGISTRATION NUMBER: 32,618  
REFERENCE/DOCKET NUMBER: 5202-D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617 498-8260  
TELEFAX: 617 876-5851  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1345 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
IMMEDIATE SOURCE:  
CLONE: HUMAN V1-1  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 138...1301  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 990...1301  
SEQUENCE DESCRIPTION: SEQ ID NO: 33:  
US-09-945-182-33

Query Match 18.2%; Score 259.2; DB 9; Length 1345;  
Best Local Similarity 57.8%; Pred. No. 2.8e-44;  
Matches 550; Conservative 0; Mismatches 378; Indels 24; Gaps 4;  
QY 464 GACAGAGATATTTGTTGATGTGTCATGCTCTACAGAAAGAGAGCTGTGGGCGG 523  
Db 364 GCAGAGCTTCTTTCGACGTGTCCAGCCTTAACGACGACGAGAGGTGGTGGTCCG 423  
QY 524 AGTTCGGCTCTTTTCGCCAGCGCCCTCAGCGCCCTTGGGGCCACAGCGCGCGCTCC 583  
Db 424 AGTTCGGCTGTCTCGCGGGGATCTCCAGAGTC--GGGCCAGGAGCTGGACTTCTCC 481  
QY 584 ACGTGAGCTCTTCCCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCG 643  
Db 482 -----GCCGTTGCTGCTGCTCCACGTGCGCGCGCGCGCGCGCGCGCGCTGC 534  
QY 644 AGGGGCGCGCGCGCGCGCTGGGAGTCTTCGACGTGTGGGAGGCGCTCGCGCCACAGC 703  
Db 535 TGTAATCGCGGCGAGCTGAGCCCTTAGTCGGTTCAGCGCTGGGAGGCGTTTCGAGTGGCG 594  
QY 704 CTTGGAAGCAGCTGTGCTTGGAGCTCGCGGCGCGCATGGGCGGAGCTGGAGCGCGGAGG 763  
Db 595 ACGCCATGAGCGCCACCGCTCGTGAACCGCGCGCGCGCGCGCGCTTCTGCTTCTGCTGC 654  
QY 764 CGAGGCGCGCGCGCGGAGACCCAGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCT 823  
Db 655 GCGAGTGGCAGCGCGCGGTGCGGAGCCGCTTGGCACTTGGCGGAGCTGGGCTTGGCTGC 714  
QY 824 TCGCGCGGAGGTGCGCGCTTCCCGAGGAGCGGCGCTTGGTGTGGTATTCACAGATCCC 883  
Db 715 CGGCGGAGGGGCTCTGCGGAGAGAGCGCGCGGTGTAGTCTCTCTCCCGACGC 774  
QY 884 AGCGCA---AGAACCTGTTTCGAGAGATGCGGAGAGCTGGGCTCGCGCGAGGCTGC-- 938  
Db 775 AGAGAAAGAGAGCTTATTCGGGAGATCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 834  
QY 939 -----GGGCCCGGGCGGGCGCGAGGGGTGTGGCGCGCGCGCGCGCGCGCGCGCC 988  
Db 835 TGGCTCTAGAGCGCTGCGCGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCTATTG 894  
QY 989 CGGATGCCAGGCGCTTGGCTGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCCCA 1048  
Db 895 CGGCGCGCAGAGAGGAGCGGCGTTGGCCGGGAGCGGCGGAGACGCGAGGCGCGCGCG 954



```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-730-772-11

Query Match      17.6%; Score 251.4; DB 10; Length 2341;
Best Local Similarity 78.7%; Pred. No. 1.2e-42;
Matches 300; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 1021 CGCGCGCGCGCGCGCGCTTCGCCAGTCCGACATGCGCAAGCGGCGCAGCAAGAGTCC 1080
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 830 CGCGAAACGCGCGCGCGCTTCGCCAGTCCGACATGCGCAAGCGGCGCAGCAAGAGTCC 889

QY 1081 AGGCTAGCTGCGAGCAAGAACGCCCTTCGACGTAATTCGAAGAGTGGGCTGGGACGAC 1140
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 890 AAGGCTGCTGCGAGTGGGAAGGACATGCTCAACTTCAAGGACATGGGCTGGGACGAC 949

QY 1141 TGGATTATCGCGCCCTTGGAGTACGAGGCTATCACTGCGAGGCTGTATCGCACTCCCG 1200
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 950 TGGATCATCGACCCCTTGAGTACGAGGCTTCCACTGCGAGGCGCTGCGAGTTCCCA 1009

QY 1201 CTGCGCTCGCACCTGGAGCCCAACACCGCCATCATCCAGACGCTGATGAATCCATG 1260
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1010 TTGCGCTCCCACTGGAGCCCAACACCGCCATCATCCAGACGCTGATGAATCCATG 1069

QY 1261 GACCCGGCTCCACCGCGCCAGCTGCTGGTCCGCCAACAAATTGACTCCCATCAGCAT 1320
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1070 GACCCGGAGTCCACAGCACCCACCTGCTGTGCGCCAGCGGCTGAGTCCCATCAGCAT 1129

QY 1321 CTATACATCGACGCGGCAATAATGTGCTACAAGCAGTACGAGCATGTGGTGGAG 1380
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1130 CTCTTATTGACTCTGCCAACAAGTGGTGTATAGCAGTATGAGGACATGGTGGTGGAG 1189

QY 1381 TCGTGGGCTGCGAGTAG 1398
      |||||  |||||  |||||
Db 1190 TCGTGGGCTGCGAGTAG 1207

RESULT 8
US-09-730-772-11
; Sequence 11, Application US/09730772
; Patent No. US2001001131A1
; GENERAL INFORMATION:
; APPLICANT: Luyten, Frank P.
; APPLICANT: Moos, Jr., Malcolm
; APPLICANT: Chang, Steven Chao-Huan
; TITLE OF INVENTION: CARTILAGE-DERIVED MORPHOGENETIC
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/730,772
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/836,081
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bartfeld, Neil S
; REGISTRATION NUMBER: 39,901
; REFERENCE/DOCKET NUMBER: NIH099.001APC
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; TELEX:
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2341 base pairs
```

```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-730-772-11

Query Match      17.6%; Score 251.4; DB 10; Length 2341;
Best Local Similarity 78.7%; Pred. No. 1.2e-42;
Matches 300; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 1021 CGCGCGCGCGCGCGCGCTTCGCCAGTCCGACATGCGCAAGCGGCGCAGCAAGAGTCC 1080
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1393 CGCGAAACGCGCGCGCGCTTCGCCAGTCCGACATGCGCAAGCGGCGCAGCAAGAGTCC 1452

QY 1081 AGGCTAGCTGCGAGCAAGAACGCCCTTCGACGTAATTCGAAGAGTGGGCTGGGACGAC 1140
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1453 AAGGCTGCTGCGAGTGGGAAGGACATGCTCAACTTCAAGGACATGGGCTGGGACGAC 1512

QY 1141 TGGATTATCGCGCCCTTGGAGTACGAGGCTATCACTGCGAGGCTGTATCGCACTCCCG 1200
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1513 TGGATCATCGACCCCTTGAGTACGAGGCTTCCACTGCGAGGCGCTGCGAGTTCCCA 1572

QY 1201 CTGCGCTCGCACCTGGAGCCCAACACCGCCATCATCCAGACGCTGATGAATCCATG 1260
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1573 TTGCGCTCCCACTGGAGCCCAACACCGCCATCATCCAGACGCTGATGAATCCATG 1632

QY 1261 GACCCGGGCTCCACCGCGCCAGCTGCTGGTCCGCCAACAAATTGACTCCCATCAGCAT 1320
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1633 GACCCGGAGTCCACACACCGCCAGCTGCTGTGCGCCAGCGGCTGAGTCCCATCAGCAT 1692

QY 1321 CTATACATCGACGCGGCAATAATGTGCTACAAGCAGTACGAGCATGTGGTGGAG 1380
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1693 CTCTTATTGACTCTGCCAACAAGTGGTGTATAGCAGTATGAGGACATGGTGGTGGAG 1752

QY 1381 TCGTGGGCTGCGAGTAGCGG 1401
      |||||  |||||  |||||  |||||
Db 1753 TCGTGGGCTGCGAGTAGCAG 1773

RESULT 9
US-09-735-849-11
; Sequence 11, Application US/09735849
; Patent No. US20010037017A1
; GENERAL INFORMATION:
; APPLICANT: Luyten, Frank P.
; APPLICANT: Moos, Jr., Malcolm
; APPLICANT: Chang, Steven Chao-Huan
; TITLE OF INVENTION: CARTILAGE-DERIVED MORPHOGENETIC
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/735,849
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/836,081
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bartfeld, Neil S
; REGISTRATION NUMBER: 39,901
; REFERENCE/DOCKET NUMBER: NIH099.001APC
```

```
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
;
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2341 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
;
; US-09-735-849-11
;
; Query Match 17.6%; Score 251.4; DB 10; Length 2341;
; Best Local Similarity 78.7%; Pred. No. 1.2e-42;
; Matches 300; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
;
; QY 1021 CGCCGGCGGGCGGCGGCTTCGCCAGTTCGCCAGTGGCAAGCGGCACGCAAGAGTCC 1080
; || || || || || || || || || || || || || || || || || || || || || ||
; Db 1393 CGCGAAAGCGGGGGCCCATCGGCCACTCGCCAGGCAAGCGCCAGCAAGACCTT 1452
;
; QY 1081 AGCTACGCTGCAGCAAGAGCCCTTCAGGAGCTTCAAGAGAGCTGGCTGGGACGAC 1140
; | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 1453 AAGGCTCGTGCAGTCCGGAAGCACTGCATGCACTTCAAGGACATGGGCTGGGACGAC 1512
;
; QY 1141 TGGATTATCGCGCCCTGGAGTACGAGGCTATCACTGCGAGGGTGTATGCGACTTCCCG 1200
; || || || || || || || || || || || || || || || || || || || || || ||
; Db 1513 TGGATCATCGACCCCTTGATGACGAGGCTTCCCATGCGAGGGCTGTCGAGTCCCA 1572
;
; QY 1201 CTGCGCTCGACCTGGAGCCCAACCAACGCGCATCATCCAGACGCTGATGAACCTCATG 1260
; || || || || || || || || || || || || || || || || || || || || || ||
; Db 1573 TTGCGCTCCACCTGGAGCCCAAGAAATATGCAATCATCCAGACCTGATGAACCTGATG 1632
;
; QY 1261 GACCCGGCTCCACCCGGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
; || || || || || || || || || || || || || || || || || || || || || ||
; Db 1633 GACCCGGAGTCCACACCCACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1692
;
; QY 1321 CTATACATCGACCGGGCAATATGTTGTTCTACAAGCAGTACGAGGACATGTTGTTGGAG 1380
; || || || || || || || || || || || || || || || || || || || || || ||
; Db 1693 CTCTTCAATGACTCTGCCAACACGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1752
;
; QY 1381 TCGTGGGCTGCGAGTAGCGG 1401
; || || || || || || || || || || || || || || || || || || || || || ||
; Db 1753 TCGTGGGCTGCGAGTAGCAG 1773
;
;
; RESULT 10
; US-08-945-459A-4
; Sequence 4, Application US/08945459A
; Patent No. US20020102633A1
;
; GENERAL INFORMATION:
; APPLICANT: MAKISHIMA, FUSAO; TAKAMATSU,
; APPLICANT: HIROYUKI; MIKI, MICHIO; KAWAI,
; APPLICANT: SHINJI; KIMURA, MICHIO; MATSUMOTO,
; APPLICANT: TOMOAKI; KATSUURA, MIEKO; ENOMOTO,
; APPLICANT: KOICHI; SATOH, YUSUKE
;
; TITLE OF INVENTION: A NOVEL PROTEIN AND
;
; NUMBER OF SEQUENCES: 4
;
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BIERMAN, MUSERLIAN AND LUCAS
; ADDRESS: LLP
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
;
; CURRENT APPLICATION DATA:
```

```
;
; APPLICATION NUMBER: US/08/945.459A
; FILING DATE: 09-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/01062
; FILING DATE: 19-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP7/322403
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP7/93664
; FILING DATE: 19-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CHARLES A. MUSERLIAN
; REGISTRATION NUMBER: 19,683
; REFERENCE/DOCKET NUMBER: 146.1275
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 661-8000
; TELEFAX: (212) 661-8002
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 357 BASE PAIRS
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: UNKNOWN
;
; US-08-945-459A-4
;
; Query Match 17.2%; Score 245.8; DB 8; Length 357;
; Best Local Similarity 81.0%; Pred. No. 1.2e-41;
; Matches 286; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
;
; QY 1043 TCGCCAGTCGCGTCAAGAGCGGCGGCAAGAGTCCAGCTACGCTCGCAGCAAGAGC 1102
; || || || || || || || || || || || || || || || || || || || || || ||
; Db 5 TCGCCACTCGCGGCAAGCGGCGGCAAGAGTCCAGCAAGAGTCCAGCTCGCAGTCCGAAGG 64
;
; QY 1103 CCTCGACGTGAACCTCAAGAGCGTGGCTGGGAGCACTGCGGCTGCGGCTGCGGCTGCGGCT 1162
; || || || || || || || || || || || || || || || || || || || || || ||
; Db 65 CACTGCATGTCACCTCAAGAGCATGGCTGGGAGCACTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 124
;
; QY 1163 AGAGGCTTACATCGGAGGGTGTATGCGACTTCCCGTGGCTGCGGCTGCGGCTGCGGCTGCGGCT 1222
; || || || || || || || || || || || || || || || || || || || || || ||
; Db 125 AGAGGCTTTCACCTGCGAGGGGCTGCGGAGTTCCTCATGCGCTCCCGCTGCGGCTGCGGCTGCGGCT 184
;
; QY 1223 CCAACCGCCCATCATCCAGACGCTGATGAACCTCCATGAGACCCCGGCTCCACCCCGCCCA 1282
; || || || || || || || || || || || || || || || || || || || || || ||
; Db 185 CGATCATCGCTCATCCAGACCCCTGATGAACCTCCATGAGACCCCGGCTCCACCCCGCCCA 244
;
; QY 1283 GCTGCTGCTGCGGCGGCAAAATGACTCCCATCAGCATTCATATACATCGAGCGGCGCAATA 1342
; || || || || || || || || || || || || || || || || || || || || || ||
; Db 245 CCGTGTGTGTCGCCAGCGGCTGAGTCCCATCAGCATTCCTTTCATGACTCTGCGCAACA 304
;
; QY 1343 ATGTGCTGTACAGCAGTACGAGGACATGCTGTTGGAGTCTGCGGCTGCGGCTGCGGCTGCGGCT 1395
; || || || || || || || || || || || || || || || || || || || || || ||
; Db 305 ACGTGTGTATAGCAGTATGAGGACATGCTGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 357
;
;
; RESULT 11
; US-09-068-253-1
; Sequence 1, Application US/09068253
; Patent No. US2002016831A1
;
; GENERAL INFORMATION:
; APPLICANT: SHIMURA, Takesada
; APPLICANT: TORIYAMA, Satsuki
;
; TITLE OF INVENTION: CARTILAGE/ BONE INDUCING MATERIALS FOR REPARATION
;
; FILE REFERENCE: 146.1286
; CURRENT APPLICATION NUMBER: US/09/068,253
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: PCT/JP96/03333
; PRIOR FILING DATE: 1996-11-14
; PRIOR APPLICATION NUMBER: JP 7/322402
; PRIOR FILING DATE: 1995-11-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
```

```

; SEQ ID NO 1
; LENGTH: 357
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(357)
; OTHER INFORMATION: Relevant amino acid residues in SEQ ID NO: 1 from
; PUBLICATION INFORMATION: 1 to 119 in WO 95/04819
; PUBLICATION INFORMATION:
; AUTHORS: HOTTEN, Gertrud
; AUTHORS: NEIDHARDT, Helge
; AUTHORS: PAULISTA, Michael
; TITLE: FAMILY
; TITLE: NEW GROWTH/DIFFERENTIATION FACTOR OF THE TGF-BETA
; PATENT FILING DATE: 1995-02-16
; RELEVANT RESIDUES: 1 TO 119
US-09-068-253-1

Query Match 17.2%; Score 245.8; DB 9; Length 357;
Best Local Similarity 81.0%; Pred. No. 1.2e-41;
Matches 286; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 1043 TCCCGATGCGCATGCGAGCGCGCAGCAAGTCCAGGCTACGCTGCGAGCAAGC 1102
DB 5 TGGCCACTCGCCAGGGCAAGCAGCAGCAAGAACCTTAAGGCTCGCTGCGAGTCGGAAG 64

QY 1103 CCTGCGACGTGAATTCAGGAGCTGGGTGGGACGACTGGATTATCGGCCCTGGAGT 1162
DB 65 CACTGCGATGTCAACTTCAGGAGCATGGGTGGGACGACTGGATCATCGCACCCCTTGAGT 124

QY 1163 ACAGGCGCTATCACTCGAGGGTGTATGCGACTTCCCGTGGCGTGGCGACCTGGAGCCCA 1222
DB 125 ACAGGCTTTCACCTCGGAGGGCTGTGCGAGTTCACATGCGCTCCACCTGGAGCCCA 184

QY 1223 CCAACACGCGCATATCCAGAGCGCTGATGAATCCATGAGACCCCGCTCCACCCGCCCA 1282
DB 185 CGAATCATGAGTCATCCAGACCGCTGATGAATCCATGAGACCCCGCTCCACACCCCA 244

QY 1283 GCTGCTGCGTGGCCACCAATTCAGTCCCATCAGCATTCATATCATGCGAGGGGCAATA 1342
DB 245 CTTGCTGTGTGCGCCAGCGGACTGAGTCCCATCAGCATCTCTTCATTGACTTGGCAACA 304

QY 1343 ATGTGCTTACAAAGCATGAGGACATGGTGTGGAGTGGTGGCGCTGCGAGT 1395
DB 305 ACCTGGGTATAGCATGATGAGGACATGGTGTGGAGTGGTGGCTGCGAGT 357

RESULT 12
US-09-945-182-29
; Sequence 29, Application US/09945182
; Patent No. US200201604941
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; Wozney, John
; Rosen, Vicki A.
; Wolfman, Neil
; Thomsen, Gerald H.
; Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSER: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/945,182
; FILING DATE: 31-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/808,324
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1203 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: murine MV1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..721
; SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-09-945-182-29

Query Match 17.2%; Score 245; DB 9; Length 1203;
Best Local Similarity 71.2%; Pred. No. 2.1e-41;
Matches 351; Conservative 0; Mismatches 137; Indels 5; Gaps 2;

QY 912 CGAGCAGCTGGGCTCGCGCAGAGCTCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 971
DB 233 CCAGGCGCTGGGTGACGAAACCAACCTGGGCGGTGCGAGGCGCGAGCGGACTGCGCTG 292

QY 972 GC-CGCGGTGCGGGCGCGCGCGGATGCCAGGCTTGGCTGCCCTCGCCCGCGCGCGCGCG 1030
DB 293 GCTGGGACTTCGGGAGNAGGAGCGGTGTGGCGCGGTGGCGCGGTGGCGCGCGCGCGCG 352

QY 1031 GGC- ---GCACGCGCTTCGCCAGTCCGCATGGCAAGCGCACGCGCAAGAGTCCAGGCTA 1086
DB 353 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 412

QY 1087 CGCTGAGCAAGAGCCCTGCGACGTGAACCTCAAGAGAGCTGGGCTGGGACGACTGGATT 1146
DB 413 CGCTGCGGTGCGCAAGTCACTGCGACGTGGACTTTAAGGAGCTGGGCTGGGACGACTG 472

QY 1147 ATCGCGCGCGCTGGAGTACGAGGCTATCACTGGGAGGGTGTATGCGACTTCCCGCTGCG 1206
DB 473 ATCGCGCGCTTAGACTACGAGGCTACCATCGGAGGGGTTCGACACTTTCCTCTGCGCG 532

QY 1207 TCGCAGCTGGAGCCCAACCAACCGCCATCATCCAGAGCGCTGATGAATCCATGGACCCC 1266
DB 533 TCGCAGCTGGAGCCTACCAACCAACCGCCATCATTCAGACGCTGCTCAACTCCATGGCG 592

QY 1267 GGCTCCACCGCGCGCGCGCTGCTGCGTGGCCCAACCAATTCAGTCCCATCAGCATTCAT 1326
DB 593 GAGCGTGGCGCGCGCGCTGCTGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 652

QY 1327 ATCGAGCGCGGCAATAATGTGGTCTACAAGCAGTACGAGGACATGGTGGTGGAGTGTGC 1386
DB 653 ATCGATGCGCGCGCAACCAACCGTGTCTACAAGCAGTACGAGGACATGGTGGTGGAGTGT 712

QY 1387 GGCTGCAGGTAGC 1399
DB 713 GGCTGCAGGTAGC 725

RESULT 13
US-09-945-182-1

```

Sequence 1, Application US/09945182  
Patent No. US2002016049A1  
GENERAL INFORMATION:  
APPLICANT: Celeste, Anthony J.  
Wozney, John  
Rosen, Vicki A.  
Wolfman, Neil  
Thomsen, Gerald H.  
Melton, Douglas A.  
TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSER: GENETICS INSTITUTE, INC.  
STREET: 87 CambridgePark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/945,182  
FILING DATE: 31-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/808,324  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Lazar, Steven R.  
REGISTRATION NUMBER: 32,618  
REFERENCE/DOCKET NUMBER: 5202-D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617 498-8260  
TELEFAX: 617 876-5851  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 926 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
CLONE: v1-1  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 571..882  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..882  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-945-182-1

Query Match 17.2%; Score 244.8; DB 9; Length 926;  
Best Local Similarity 64.7%; Pred. No. 2.2e-41;  
Matches 423; Conservative 0; Mismatches 207; Indels 24; Gaps 3;  
QY 771 GCGCGCGGGGACCCAGCAACCGCGCCCGGACCTGCGGAGTCTGGGCTTCGGCG 830  
Db 234 GCGCGAGTGGCGCGCGCGGTCCGAGCCGTTGGCACTGCGGCGACTGGGCTTCGGCTG 293  
QY 831 G-----AGGTGCGGCCCTCCCGAGAGCGGCGCTGCTGGTGGTATTCACGAGTC 881  
Db 294 GCGGCGGCGAGGGGCTCTGCGGCGAGAGCGCGGCTAGTCTCTCTCCCGCGAC 353  
QY 882 CCAGCGCA---AGACCTGTTCCAGAGATGCGGAGAGCTGGGCTCGGCGAGGCTGC 938  
Db 354 GCAGAGAAAGAGAGCTTATTCGGGAGATCCGCGCCAGCGCCGCGCGCTCGGGCGCGC 413

QY 939 -----GGCGCGGGCGCGGGCGCGCGAGGGGCTGTGGCGCGCGCGCTCGGGCGC 986  
Db 414 TCTGGCCTCAGAGCGCTGCCCGACCCAGGAACCGGACCGCGTCCGCAAGGCGAGTCAT 473  
QY 987 CCCGATGCCAGGCTTGGCTGCCCTCGCGCGCGCGCGCGCGCGCGCGCGCTTCGC 1046  
Db 474 TGGCGCGCGCAGCAGCGAGGAGGAGCGCGTGGCGCGGACGCGACAGCGAGCGGAGCGG 533  
QY 1047 CAGTGCCTATGCCAAGCGCGCAGGCAAGAGTCCAGGCTACCTGCAGCAAGAGCCCT 1106  
Db 534 CGGGGCGCGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCAGCGCAAGCGGT 593  
QY 1107 GCACGTGAACCTCAAGGAGCTGGCTGGGACGACTGGATTATCGGCGCGCGCTGGAGTAGCA 1166  
Db 594 GCACGTGGACTTCAAGGAGCTGGCTGGGACGACTGGATCATCGGCGCGCTGGAGTAGCA 653  
QY 1167 GGCCTATCACTCGAGGGGTGATGCGACTTCCCGTGGCTGCGACCTGGAGGCCACCAA 1226  
Db 654 GCGGTACCACTCGAGGGGCTTGGGACTTCCTTTGGTTCGCACTCGAGGCCACCAA 713  
QY 1227 CCACGCCATCATCCAGAGCGCTGATCACTCCATCAGATCCAGCCCGCGCTCCACCGCGCGCTG 1286  
Db 714 CCATGCCATCATTCAGAGCGCTGCTCAACTCCATGGCAACAGCGCGCGCGCTCCCTG 773  
QY 1287 CTGCGTGGCCACCAATGACTCCCATCAGCATCTTATACATCGACGCGGCGCAATAATGT 1346  
Db 774 CTGTGTGCCAGCGCGCTCAGCGCCCATCAGCATCTTACATCGAGCGCGCGCAACAGCT 833  
QY 1347 GGTCTACAAGCAGTACGAGGACATGGTGGTGGAGTCTGTGGGCTCGAGGTAGCG 1400  
Db 834 TGTCTACAAGCAATACGAGGACATGGTGGTGGAGGCTGCGGCTCGAGGTAGCG 887

RESULT 14  
US-09-880-708-9  
Sequence 9, Application US/09880708  
Patent No. US20020165361A1  
GENERAL INFORMATION:  
APPLICANT: Lee, Se-Jin  
Huynh, Thanh  
TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-5  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Gray Cary Ware & Freidenrich LLP  
STREET: 4365 Executive Drive, Suite 1600  
CITY: San Diego  
STATE: CA  
COUNTRY: USA  
ZIP: 92121-2189  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/880,708  
FILING DATE: 12-Jun-2001  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/145,060  
FILING DATE: <Unknown>  
APPLICATION NUMBER: 08/003,144  
FILING DATE: 12-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Lisa A. Halle, Ph.D.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/057002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 858/677-1456  
TELEFAX: 619/677-1465  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2329 base pairs  
TYPE: nucleic acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; IMMEDIATE SOURCE:
; CLONE: GD-5
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 322...1806
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-880-708-9

Query Match          16.8%;   Score 239;   DB 9;   Length 2329;
Best Local Similarity 74.9%;   Pred. No. 3.8e-40;
Matches 299;   Conservative 0;   Mismatches 100;   Indels 0;   Gaps 0;

Qy  1021 CGCCGGGGGGCGGCGACGGCCCTTCGCCAGTGCCTCCATGGCAAGCGCGACGGCAAGAAGTCC 1080
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  1432 CGGCGGAACGCCGGGCCCATTTGCCCAATGCCAGGCGACGCCAGCAGAACCTC 1491

Qy  1081 AGGCTAGCTGCAGCAGAGAAGCCCTTGCAGCTGAACCTTCAAGGAGCTGGGTGGGACGAC 1140
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  1492 AAGGCTCGCTCAGTCGCAAGCCCTTGCAATGCAACTTCAAGGACATGGCTGGGACGAC 1551

Qy  1141 TGGATTATCGGCCCTTGGAGTAGAGGCCCTATCACTCGGAGGGTGTATCGCACTTCCCG 1200
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  1352 TGGATATCGCACCTCTTGAGTATGAGGCCCTTCCACTGCGAAGGACTGTGTAGTTCCCC 1611

Qy  1201 CTGGCTCGCACTTGGAGCCCCACCAACACGCCCATCTTCCAGAGCGCTCATGAACCTCCATG 1260
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  1612 TTGGCTGCCACTTGGAGCCCCAACAACACCGATCATTCAGACCTTAATGACTCTATG 1671

Qy  1261 GACCCGGGCTCCACCCCGCCGAGCTGTGCTGCCACCAAAATTGACTCCCATCAGCAATT 1320
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  1672 GACCTGAATCCACACACCCACTTGTGTGCTTACACGGCTGAGTCTATTAGCATC 1731

Qy  1321 CTATACATCAGCGGGGCAATAATGTGGTCTACAGAGAGTACGAGGACATGGTGGTGGAG 1380
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  1732 CTCCTCATCGACTCTGCCAACAGCTGGTGATAAACAGTACGAGGACATGTCGTGGAA 1791

Qy  1381 TCGTGGCGCTCGAGGTAGCGGTGCTTTCCCGCGCGCTT 1419
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  1792 TCTTGTGGCTGCAAGTAGACACACCGGGCCCACTGTCTT 1830

RESULT 15
US-09-833-381-582
; Sequence 582, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 582
; LENGTH: 475
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-582

Query Match          16.6%;   Score 237;   DB 10;   Length 475;
Best Local Similarity 74.1%;   Pred. No. 7.5e-40;
Matches 300;   Conservative 0;   Mismatches 105;   Indels 0;   Gaps 0;

Qy  996 CAGCGCTTGGCTCGCCCTCGCCCGCGCGGGCGGCGACGCGCTTCGCCAGTTCGCGCA 1055
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  8 CAGAGGAGGAGGACGGCTTGGCGGGACCGCGACATCGCAGGCGACGGCGGGGGCGC 67

Qy  1056 TGGCAAGCGGCACGGCAAGAAGTCCAGGCTACGCTGCAGCAAGAAGCCCTGCACGTGAA 1115

```





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2002, 01:08:25 ; Search time 82 Seconds  
(without alignments)  
739,378 Million cell updates/sec

Title: US-09-825-751A-20

Perfect score: 2447

Sequence: 1 MDTPRVLLSAVFLISFLMDL.....GNVVYKQYEDMVVESCGR 455

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2447	100.0	455	22	AAW50216 Human growth/diffe
2	2447	100.0	455	23	AAU79173 Human growth/diffe
3	2437	99.6	455	23	AAE17604 Human extracellular
4	1798	73.5	436	17	AAW95636 Cartilage-derived
5	1747	71.4	321	16	AAW78730 Human mature VL-1
6	1747	71.4	321	18	AAW26591 Human bone morphog
7	1747	71.4	321	22	AAE10982 Human full length
8	1207.5	49.3	263	16	AAW78739 Murine mv2 protein
9	1207.5	49.3	263	18	AAW26595 Murine BMP-13 homo
10	1202.5	49.1	263	22	AAE10985 Murine partial mv2

11	927.5	37.9	501	16	AAW69600 New TGF-beta famil
12	927.5	37.9	501	18	AAW36100 Human MP52. Homo
13	927.5	37.9	501	18	AAW19210 Human TGF-beta pro
14	927.5	37.9	501	18	AAW11900 Human high mol. wt
15	927.5	37.9	501	18	AAW01799 Human MP52 protein
16	927.5	37.9	501	18	AAW12770 Human bone morphog
17	927.5	37.9	501	19	AAW44868 TGF-beta superfam1
18	927.5	37.9	501	19	AAW33008 Human MP52. Homo
19	919.5	37.6	495	15	AAW60022 Growth differentia
20	919.5	37.6	495	22	AAW84550 Amino acid sequenc
21	918.5	37.5	501	17	AAW95635 Cartilage-derived
22	918.5	37.5	501	21	AAW92034 Human growth diffe
23	917.5	37.5	401	14	AAW40800 TGF-beta-like clon
24	916.5	37.5	501	22	AAW70529 Human TGF-beta MP5
25	765	31.3	388	16	AAW78734 Human bone morphog
26	765	31.3	388	18	AAW26592 Human bone morphog
27	765	31.3	388	22	AAE10986 Human bone morphog
28	731	29.9	134	16	AAW66867 GDF-6. Mus sp. A
29	731	29.9	134	21	AAW12986 Murine growth diff
30	691.5	28.3	294	16	AAW78729 Human bone morphog
31	691.5	28.3	294	18	AAW26589 Human bone morphog
32	691.5	28.3	294	22	AAE10972 Human bone morphog
33	666	27.2	120	23	AAW51933 Murine TGFbeta pro
34	627	25.6	411	16	AAW78740 Fusion of BMP-2 pr
35	627	25.6	411	18	AAW26597 BMP-2 propeptide/B
36	627	25.6	411	22	AAE10983 Human BMP2 propept
37	618	25.3	240	16	AAW78738 Murine mv1 protein
38	618	25.3	240	18	AAW26594 Murine BMP-12 homo
39	618	25.3	240	22	AAE10984 Murine partial mv1
40	601	24.6	184	22	ABG29375 Novel human diagno
41	575	23.5	102	21	AAW09554 Human CDMP-2/GDF-6
42	575	23.5	102	21	AAW02820 Human CDMP-2/GDF-6
43	575	23.5	102	21	AAW92579 CDMP-2/GDF-6 finge
44	572	23.4	102	21	AAW09555 Murine GDF-6, SEQ
45	572	23.4	102	21	AAW02821 Mouse GDF-6 amino

ALIGNMENTS

RESULT 1				
AAW50216				
ID	AAW50216 standard; Protein: 455 AA.			
AC	AAW50216;			
XX				
DT	07-JAN-2002 (first entry)			
XX				
DE	Human growth/differentiation factor-6-like protein AMF10.			
XX				
KW	AMF10; human; growth/differentiation factor-6; cancer;			
XX				
KW	cell proliferation; astrocytoma; glioma; therapy; diagnosis.			
XX				
OS	Homo sapiens.			
XX				
FX	Key Location/Qualifiers			
FT	Peptide 1..22			
FT	Protein /label= Signal_peptide			
FT	23..455 /label= Mature_protein			
XX				
PN	WO200174897-A2.			
XX				
PD	11-OCT-2001.			
XX				
PF	03-APR-2001; 2001WO-US10892.			
XX				
PR	03-APR-2000; 2000US-194314P.			
XX				
PA	16-AUG-2000; 2000US-225693P.			
XX				
XX	(CURA-) CURAGEN CORP.			
XX				
PI	Vernet CAM, Burgess CE, Fernandes E, Taupier RJ, Quinn KE;			

PI Spytek KA, Rastelli L, Herrmann JL;  
XX WPI: 2001-626395/72.  
DR N-PSDB; AAI70203.  
XX New AMF1-10 polypeptides and encoding polynucleotides, useful for  
PT treating or preventing disorders related to modulation of cell  
PT movement, cell signal processing, cell adhesion or migration pathways  
PT e.g., cancer  
XX  
XX Claim 1; Page 44-45; 134pp; English.  
PS  
XX  
XX The present sequence is that of the novel, secreted human  
CC growth/differentiation factor-6 (GDF6)-like protein, AMF10.  
CC AMF10 is expressed in astrocytoma and glioma-derived tissue. DNA  
CC encoding the AMF8 may be useful in gene therapy, and the protein  
CC may also be used as a therapeutic, especially in treatment of  
CC cancer and other cell proliferative disorders. Generally, the  
CC AMF1-10 (AMFX) nucleic acids and proteins of the invention are useful  
CC for treating or preventing AMFX-associated disorders, e.g. a disorder  
CC related to cell signal processing and metabolic pathway modulation,  
CC cell adhesion or migration pathway modulation, chemoresistance,  
CC radiotherapy resistance, survival in trophic factor limited  
CC secondary tissue site microenvironments, connective tissue  
CC disorders, tissue remodeling, oncogenesis, cancer of the breast,  
CC ovary, cervix, prostate, endometrium, stomach, colon, lung,  
CC bladder, kidney, brain, and soft-tissue, cellular transformation,  
CC developmental tissue remodeling, inflammation, blood clot  
CC formation and resorption, haematopoiesis, angiogenesis, multidrug  
CC resistance related to organic anion transporters, malignant disease  
CC progression, autocrine and paracrine regulation of cell growth, and  
CC cellular responses to external stimuli, and other diseases,  
CC disorders, etc. (all claimed). AMFX proteins are also used for  
CC screening drugs or compounds that modulate AMFX protein activity or  
CC expression as well as to treat disorders characterized by  
CC insufficient or excessive production of AMFX protein.  
XX  
SQ Sequence 455 AA;  
Query Match 100.0%; Score 2447; DB 22; Length 455;  
Best Local Similarity 100.0%; Pred. No. 1.2e-199;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDTPRVLLSAVFLISFLMDLPFGQOASISSSCSSAELGSTKGRSKGKMQRAPRDSDA 60  
DB 1 MDTPRVLLSAVFLISFLMDLPFGQOASISSSCSSAELGSTKGRSKGKMQRAPRDSDA 60  
QY 61 GREGQEPQRPQDEPRAQPPRAQPPGPRVVPVPEHYMLSIYRTYSIAEKLGINASFQSS 120  
DB 61 GREGQEPQRPQDEPRAQPPRAQPPGPRVVPVPEHYMLSIYRTYSIAEKLGINASFQSS 120  
QY 121 SKSANTITSFVDRGLDLSHTPLRQKYLFDVSMLSKDEELVGAELRFRQAPSAPWGPP 180  
DB 121 SKSANTITSFVDRGLDLSHTPLRQKYLFDVSMLSKDEELVGAELRFRQAPSAPWGPP 180  
QY 181 AGPLHVQLFPCLSPDLLDARTLDPOGAPAGWEVDVWQGLRHQPWKOLCLELRAAWGEL 240  
DB 181 AGPLHVQLFPCLSPDLLDARTLDPOGAPAGWEVDVWQGLRHQPWKOLCLELRAAWGEL 240  
QY 241 DAGEAEARAGPQQPPPLDLSLGLGRRVRPPQERALLVFTTRSQKNLFAEMREQLGSA 300  
DB 241 DAGEAEARAGPQQPPPLDLSLGLGRRVRPPQERALLVFTTRSQKNLFAEMREQLGSA 300  
QY 301 EAAGGAGAGSWPPSPGAPDAPRPLPSFGRRRRRTAFASRHKRGKSLRCSKKPLH 360  
DB 301 EAAGGAGAGSWPPSPGAPDAPRPLPSFGRRRRRTAFASRHKRGKSLRCSKKPLH 360  
QY 361 VNFKELGWDWIIAPLEYAEYHCEGVCDFPLRSLHLEPTNHHIIOTLMSMPDGGSTPPSCC 420  
DB 361 VNFKELGWDWIIAPLEYAEYHCEGVCDFPLRSLHLEPTNHHIIOTLMSMPDGGSTPPSCC 420  
QY 421 VPTKLTPTISILYIDAGNNVVYKQYEDMVVESCGR 455  
DB 421 VPTKLTPTISILYIDAGNNVVYKQYEDMVVESCGR 455

Db 421 VPTKLTPTISILYIDAGNNVVYKQYEDMVVESCGR 455  
RESULT 2  
AAU79173  
ID AAU79173 standard; Protein; 455 AA.  
XX AC AAU79173;  
XX 02-JUL-2002 (first entry)  
XX Human growth/differentiation factor 6-like protein NOV9.  
DE Human; growth/differentiation factor 6-like protein; NOV9;  
XX KW vaccine; cancer; tumour; bone disorder; avascular necrosis; allergy;  
XX KW haematopoietic disorder; immune disorder; endometriosis; renal disease;  
XX KW infection; inflammatory disease; lung disease; scleroderma; ataxia;  
XX KW bowel disease; appendicitis; blood disorder; cardiovascular disorder;  
XX KW graft versus host disease; GVHD; lymphoedema; brain disorder;  
XX KW ocular disorder; hepatitis C virus infection; cardiac disorder;  
XX KW autosomal dominant deafness; DFNA-2.  
OS Homo sapiens.  
XX WO200214368-A2.  
PN 21-FEB-2002.  
PD 16-AUG-2001; 2001WO-US25624.  
XX PF 16-AUG-2000; 2000US-225692P.  
XX PR 16-AUG-2000; 2000US-225693P.  
XX PR 16-AUG-2000; 2000US-225837P.  
XX PR 18-AUG-2000; 2000US-226236P.  
XX PR 18-AUG-2000; 2000US-226353P.  
XX PR 22-AUG-2000; 2000US-227085P.  
XX PR 23-AUG-2000; 2000US-227395P.  
XX PR 24-AUG-2000; 2000US-227492P.  
XX PR 24-AUG-2000; 2000US-227600P.  
XX PR 14-MAR-2001; 2001US-275952P.  
XX (CURA-) CURAGEN CORP.  
PA Zerhusen BD, Padigaru M, Spytek KA, Spaderna SK, Gangolli EA;  
PI Rastelli L, Burgess CE, Majumder K, Shimkets R, Mishra V;  
PI Vernet CAM, Szekeres ES, Grosse WM, Alsobrook JP, Liu X;  
PI Gerlach VL, Ellerman K, Smithson G, Peyman J, Stone D;  
PI Macdougall J;  
XX WPI: 2002-329571/36.  
DR N-PSDB; ABK48393.  
XX Novel cytoplasmic, nuclear membrane bound and secreted NOVX  
PT polypeptides, useful for treating cancers and tumours, bone disorders,  
PT Paget's disease, haematopoietic disorders, spinal diseases and immune  
PT disorders  
XX Claim 1; Page 86; 234pp; English.  
XX The present invention relates to new isolated NOVX polypeptides named  
CC NOV1-NOV9. The invention can be used for identifying an agent (a cellular  
CC receptor or downstream effector) that binds to the polypeptide. The  
CC molecules of the invention are useful for treating or preventing  
CC NOVX-associated disorders in humans. The antibody of the invention is  
CC useful for determining the presence or amount of NOVX in a sample, and  
CC for treating a pathological state in a mammal. The method of the  
CC invention is useful for determining the presence of an amount of NOVX in  
CC a sample which is used as a marker for cancerous cell or tissue type.  
CC The molecules of the invention are useful in the manufacture of a  
CC medicament for treating or preventing cancer, tumour, bone disorders,  
CC avascular necrosis, allergy, haematopoietic disorders, immune disorders,  
CC endometriosis, renal diseases, infections, inflammatory diseases, lung  
CC diseases, scleroderma, ataxia, bowel diseases, appendicitis, blood

CC disorders, cardiovascular disorders, graft versus host disease (GVHD),  
CC lymphoedema, brain disorders, ocular disorders, hepatitis C virus  
CC infection, cardiac disorders and autosomal dominant deafness (DFNA-2).  
CC The present amino acid sequence represents the human  
CC growth/differentiation factor 6-like protein NOV9 that is one of the  
CC NOVX proteins described in the invention.  
XX  
SQ Sequence 455 AA;

Query Match 100.0%; Score 2447; DB 23; Length 455;  
Best Local Similarity 100.0%; Pred. No. 1.2e-199;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDTPRVLLSAVFLISFLWDLPGFQQASISSSCSSAELGSTKGMRSKEGKMQRAPRDSDA 60  
DB 1 MDTPRVLLSAVFLISFLWDLPGFQQASISSSCSSAELGSTKGMRSKEGKMQRAPRDSDA 60  
QY 61 GREGQEPQPRQDEPRAQPPRAQEPFGPRVVPVHEYMLSIYRTYSIAEKLGINASFFQS 120  
DB 61 GREGQEPQPRQDEPRAQPPRAQEPFGPRVVPVHEYMLSIYRTYSIAEKLGINASFFQS 120  
QY 121 SKSANTITSFVDRGLDDLSHTPLRRQKYLFDVSMLSKDELVGAELRFPQAPSAPWGPP 180  
DB 121 SKSANTITSFVDRGLDDLSHTPLRRQKYLFDVSMLSKDELVGAELRFPQAPSAPWGPP 180  
QY 181 AGPLHVQLFPCLSPLLDARTLDPQAGPAGWEVFDVWQGLRHQPWKQLCLELRAAWGEL 240  
DB 181 AGPLHVQLFPCLSPLLDARTLDPQAGPAGWEVFDVWQGLRHQPWKQLCLELRAAWGEL 240  
QY 241 DAGEAEARAGPQQPPDLRLSGFGRRVPPQERALLVVFTRSQKRNLFPAEMREQLGS 300  
DB 241 DAGEAEARAGPQQPPDLRLSGFGRRVPPQERALLVVFTRSQKRNLFPAEMREQLGS 300  
QY 301 EAAGPAGAGEGSWPPSPGADARPWLPSPGRRRRRTAFASRHKRGKSKKPLH 360  
DB 301 EAAGPAGAGEGSWPPSPGADARPWLPSPGRRRRRTAFASRHKRGKSKKPLH 360  
QY 361 VNFKELGWDDWIIAPLEYEAYHCEGVCDPLRSHLEPTNHAIQTLMNSMDPGSTPPSCC 420  
DB 361 VNFKELGWDDWIIAPLEYEAYHCEGVCDPLRSHLEPTNHAIQTLMNSMDPGSTPPSCC 420  
QY 421 VPTKLTPISTLYIDAGNNVYKQYEDMVVESCGR 455  
DB 421 VPTKLTPISTLYIDAGNNVYKQYEDMVVESCGR 455

RESULT 3  
ID AAE17604 standard; Protein; 455 AA.  
AC AAE17604;  
XX

22-APR-2002 (first entry)

Human extracellular messenger (XMB5)-6 protein.

Human; extracellular messenger; neurological disorder; epilepsy; XMB5-6;  
KW Alzheimer's disease; autoimmune disorder; renal tubular acidosis; stroke;  
KW acquired immune deficiency syndrome; AIDS; Addison's disease; cytostatic;  
KW developmental disorder; anaemia; Cushing's syndrome; endocrine disorder;  
KW vascular malformation; cell proliferative disorder; gene therapy; cancer;  
KW neuroprotective; trauma; hypopituitarism; hypothyroidism; antihelminthic;  
KW hyperthyroidism; gonadal steroid hormone; pancreatic disorder; nontropic;  
KW diabetes mellitus; immunosuppressive; anti-inflammatory; antibacterial;  
KW antiviral; antifungal; parasitic; protozoal; allergy.

OS Homo sapiens.

PN WO200194587-A2.

XX 13-DEC-2001.

PF 06-JUN-2001; 2001WO-US18476.

XX 06-JUN-2000; 2000US-210233P.  
PR 23-JUN-2000; 2000US-213465P.  
PR 14-NOV-2000; 2000US-249019P.  
XX (INCY-) INCYTE GENOMICS INC.  
XX Lal P, Yue H, He A, Nguyen DB, Wallia N, Gandhi AR, Azimzai Y;  
PI Bandman O, Tang YT, Lu Y, Baughn MR, Duggan BM, Lee S, Hafalia A;  
PI Policky JL;  
XX WPI: 2002-154573/20.  
DR N-PSDB; AAD28496.  
XX New extracellular messenger polypeptides and polynucleotides encoding  
PT them, useful for diagnosing, treating or preventing e.g. neurological,  
PT autoimmune, inflammatory, developmental and endocrine disorders  
XX Claim 1; Page 113-114; 123pp; English.  
XX The present invention relates to new extracellular messenger polypeptides  
CC and polynucleotides encoding them. XMB5 is useful in the diagnosis,  
CC treatment and prevention of neurological disorders (e.g. epilepsy,  
CC stroke, or Alzheimer's disease), autoimmune/inflammatory disorders (e.g.  
CC acquired immune deficiency syndrome, AIDS, Addison's disease, or  
CC allergies), developmental disorders (e.g. renal tubular acidosis, anaemia  
CC or Cushing's syndrome), endocrine disorders (e.g. hypophysectomy,  
CC aneurysm or vascular malformation), and cell proliferative disorders  
CC (e.g. cancer), and in the assessment of the effects of exogenous  
CC compounds on the expression of nucleic acid and amino acid sequences of  
CC extracellular messengers. XMB5 may also be used in the treatment of  
CC viral, bacterial, fungal, parasitic, protozoal and helminthic  
CC infections, trauma, disorders associated with hypopituitarism,  
CC hypothyroidism, hyperthyroidism or gonadal steroid hormones, and  
CC pancreatic disorders such as type I or type II diabetes mellitus. The  
CC XMB5 may be used for somatic or germline gene therapy. The present  
CC sequence is human XMB5-6 protein.  
XX Sequence 455 AA;  
SQ

Query Match 99.6%; Score 2437; DB 23; Length 455;  
Best Local Similarity 99.8%; Pred. No. 8.8e-199;  
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDTPRVLLSAVFLISFLWDLPGFQQASISSSCSSAELGSTKGMRSKEGKMQRAPRDSDA 60  
DB 1 MDTPRVLLSAVFLISFLWDLPGFQQASISSSCSSAELGSTKGMRSKEGKMQRAPRDSDA 60  
QY 61 GREGQEPQPRQDEPRAQPPRAQEPFGPRVVPVHEYMLSIYRTYSIAEKLGINASFFQS 120  
DB 61 GREGQEPQPRQDEPRAQPPRAQEPFGPRVVPVHEYMLSIYRTYSIAEKLGINASFFQS 120  
QY 121 SKSANTITSFVDRGLDDLSHTPLRRQKYLFDVSMLSKDELVGAELRFPQAPSAPWGPP 180  
DB 121 SKSANTITSFVDRGLDDLSHTPLRRQKYLFDVSMLSKDELVGAELRFPQAPSAPWGPP 180  
QY 181 AGPLHVQLFPCLSPLLDARTLDPQAGPAGWEVFDVWQGLRHQPWKQLCLELRAAWGEL 240  
DB 181 AGPLHVQLFPCLSPLLDARTLDPQAGPAGWEVFDVWQGLRHQPWKQLCLELRAAWGEL 240  
QY 241 DAGEAEARAGPQQPPDLRLSGFGRRVPPQERALLVVFTRSQKRNLFPAEMREQLGS 300  
DB 241 DAGEAEARAGPQQPPDLRLSGFGRRVPPQERALLVVFTRSQKRNLFPAEMREQLGS 300  
QY 301 EAAGPAGAGEGSWPPSPGADARPWLPSPGRRRRRTAFASRHKRGKSKKPLH 360  
DB 301 EAAGPAGAGEGSWPPSPGADARPWLPSPGRRRRRTAFASRHKRGKSKKPLH 360  
QY 361 VNFKELGWDDWIIAPLEYEAYHCEGVCDPLRSHLEPTNHAIQTLMNSMDPGSTPPSCC 420  
DB 361 VNFKELGWDDWIIAPLEYEAYHCEGVCDPLRSHLEPTNHAIQTLMNSMDPGSTPPSCC 420  
QY 421 VPTKLTPISTLYIDAGNNVYKQYEDMVVESCGR 455

Db 421 VPTKLTPIISLYIDAGNNVYKQYEDMVVSGCR 455

RESULT 4

AAR95636  
ID AAR95636 standard; Protein; 436 AA.

AC AAR95636;

DT 25-OCT-1996 (first entry)

XX Cartilage-derived morphogenetic protein-2.

KW Cattle; cartilage-derived morphogenetic protein-2; CDMP-2;  
articular cartilage; chondrogenic; vulnery; implantation;  
chondromalacia; osteoarthritis; therapy; joint repair.

OS Bos taurus.

XX Key Location/Qualifiers

FT Region 1..312

FT /note= "Pro-region"

FT Modified-site 89..91

FT /note= "N-glycosylation site"

FT Cleavage-site 313..316

FT /note= "Proteolytic processing site"

FT Domain 317..436

FT /note= "C-terminal mature domain"

FT Peptide 352..382

FT /note= "Consensus conserved motif (AAR95641)"

XX

PN W09614335-A1.

XX

PD 17-MAY-1996.

XX

PF 07-NOV-1994; 94WO-US12814.

XX

PR 07-NOV-1994; 94WO-US12814.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Chang SC, Luyten FP, Moos M;

XX

DR WPI; 1996-251714/25.

XX

DR N-PSDB; AAT31602.

XX

PT New purified cartilage extracts and proteins - used to stimulate the  
development and repair of cartilage in vivo.

XX

PS Claim 11; Fig 2; 34pp; English.

XX

CC The sequence represents cattle articular cartilage-derived  
morphogenetic protein-2 (CDMP-2). The N-terminal methionine and  
signal peptide is missing, but part of the pro-region, a typical  
proteolytic cleavage site and a C-terminal domain containing 7  
highly conserved Cys residues characteristic of the transforming  
growth factor-beta gene family are present. A single N-glycosylation  
site is located in the pro-region. A consensus highly conserved  
motif in CDMP proteins (AAR95641) is present in the C-terminal domain.  
CDMP-2 is present in a purified cartilage extract (claimed) which  
stimulates local cartilage formation and repair when combined with a  
matrix and implanted in a mammal. The protein may be used in therapy  
of e.g. chondromalacia or osteoarthritis, to heal joint surfaces, or  
to repair cartilage after reconstructive surgery.

XX Sequence 436 AA;

Query Match 73.5%; Score 1798; DB 17; Length 436;  
Best Local Similarity 80.8%; Pred. No. 1.8e-144;  
Matches 354; Conservative 18; Mismatches 48; Indels 18; Gaps 8;

Qy 33 SSAELGSTKGMRSRKEGKMRAPRDSAGRE---GQEPQRPQDEPR---AQQRAQEP 86

Db 2 ASAEELGSAKMRTRKEGRMPRAPRENATAREPLDRQEPPEPPQEQRRPPQPEAREPP 61  
Qy 87 GRGPRVVPHEYMLSYRTYSIAEKLGINASFQSSKSANTITSFYDRGLDLSHTPLRQ 146  
Db 62 GRGPRVVPHEYMLSYRTYSIAEKLGINASFQSSKSANTITSFYDRGLDLSHTPLRQ 121  
Qy 147 KYLFDVMSLSDKEELVGAELRFLROAPSAPWGPAGPLHVQLFPCLSPDLLDARTLDPOG 206  
Db 122 KYLFDVSTLSDKEELVGADVRLFRQAPALAPAAAPLAALRLP-VAPAAGSAEP-GPAG 179  
Qy 207 APPAGWEVDFVWQGLRHQPKQLCLELRAAW-GELDAGEAEARARQPQPPPDLSLGF 265  
Db 180 APRPGWEVDFVWQGLRHQPKQLCLELRAAWGEGPAAEAEARTPGQPQPPPDLSLGF 239  
Qy 266 GRVRPPQPERALLVFTRSQRNLFAEMREQLGSA-EAAGPGAGAGSGW-----PPPS 317  
Db 240 GRVRTPOERALLVFTRSQRNLFAEMREQLGSAEYVYVGGGAGSGGPPPPPPPPPS 299  
Qy 318 GAPDARPLPSPGRRRRRTAFASRHGKRGKSRKRLCSKKPLHVNFKELGWDDWIITAPLE 377  
Db 300 GTPDAGLWSPG-RRRTAFASRHGKRGKSRKRLCSKKPLHVNFKELGWDDWIITAPLE 358  
Qy 378 YEAYHCEGVCDFFPLRSHLEPTNHAIITQTLNMSMDPGSTPPSCVPTKLTPTISLYIDAGN 437  
Db 359 YEAYHCEGVCDFFPLRSHLEPTNHAIITQTLNMSMDPGSTPPSCVPTKLTPTISLYIDAGN 418  
Qy 438 NVYKQYEDMVVSGCR 455  
Db 419 NVYNEYEEMVYVSGCR 436  
RESULT 5  
AAR78730  
ID AAR78730 standard; Protein; 321 AA.  
XX AC AAR78730;  
XX DT 30-NOV-1995 (first entry)  
XX DE Human mature VL-1 (BMP-13) encoding sequence.  
XX KW Bone morphogenetic protein; VL-1; tendon; ligament; tendinitis.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
FT Peptide 1..201  
FT Protein 202..321  
FT /label= mature protein  
FT Misc-difference 202..321  
FT /note= "claimed"  
FT Misc-difference 220..321  
FT /note= "claimed"  
XX W09516035-A.  
XX PD 15-JUN-1995.  
XX PF 06-DEC-1994; 94WO-US14030.  
XX PR 02-NOV-1994; 94US-0333576.  
XX PR 07-DEC-1993; 93US-0164103.  
XX PR 25-MAR-1994; 94US-0217780.  
XX (GEM ) GENETICS INST INC.  
XX PA (HARD ) HARVARD COLLEGE.  
XX Celeste AJ, Melton DA, Rosen VA, Thomsen GH, Wolfman NM;  
Wozney JW;  
DR WPI; 1995-224320/29.  
DR N-PSDB; AAQ96208.

XX Bone morphogenetic proteins -12 and -13 and corresp. DNA - used in  
PT compsn. for inducing tendon/ligament-like tissue formation  
XX  
XX  
PS Claim 14; Page 62-64; 84pp; English.  
XX  
XX BMP-12 related proteins are a subset of the BMP/TGF-beta/Vg-1  
CC family of proteins, including BMP-12 and Vt-1. Vt-1 is designated  
CC BMP-13. Like BMP-12, it is expected that BMP-13, as expressed by  
CC mammalian cells such as CHO cells, exists as a heterogeneous popn.  
CC of active species of BMP-13 protein with varying N-termini. It is  
CC expected that all active species will contain the AA sequence  
CC beginning with the 19th Cys residue of the mature protein until  
CC the 119th residue of the mature protein or until the stop codon  
CC after the 120th residue of the mature protein. Other active  
CC species contain additional AA sequence in the N-terminal direction.  
CC AAQ96208 is a partial DNA sequence and AAR78730 is the derived AA  
CC sequence of a portion of the 2.5 kb DNA insert of the plasmid  
CC subclone pGEMJLDC31/2.5, derived from clone lambdaJLDC31.  
XX  
SQ Sequence 321 AA;  
  
Query Match 71.4%; Score 1747; DB 16; Length 321;  
Best Local Similarity 100.0%; Pred. No. 2.7e-140;  
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 137 DLSHTPLRRQKYLFDVSMLSKDELVGAEELRLFRQAPSPAPGPPAGPLHVQLFPCLSPLL 196  
Db 3 DLSHTPLRRQKYLFDVSMLSKDELVGAEELRLFRQAPSPAPGPPAGPLHVQLFPCLSPLL 62  
  
QY 197 LDARTLDPQAGPAGWEVDFVWQGLRHQPKQLCLELRAANGELDAGEAEARAGPQQPP 256  
Db 63 LDARTLDPQAGPAGWEVDFVWQGLRHQPKQLCLELRAANGELDAGEAEARAGPQQPP 122  
  
QY 257 PPDLSLGFGRVRPPQERALLVVFTRSORKNLFAEMREQLGSAEAGPGAGAGSWPPP 316  
Db 123 PPDLSLGFGRVRPPQERALLVVFTRSORKNLFAEMREQLGSAEAGPGAGAGSWPPP 182  
  
QY 317 SGAPDARPLWPSGGRRRRTAFASRHGKRHGKKSRLRCSKKPLHVNFKELGWDWIIAPL 376  
Db 183 SGAPDARPLWPSGGRRRRTAFASRHGKRHGKKSRLRCSKKPLHVNFKELGWDWIIAPL 242  
  
QY 377 EYEAHCEGVCDFFLRSHLEPTNHAIIQTLMNSMDPGSTPPSCCVPTKLTPIISILIYIDAG 436  
Db 243 EYEAHCEGVCDFFLRSHLEPTNHAIIQTLMNSMDPGSTPPSCCVPTKLTPIISILIYIDAG 302  
  
QY 437 NNVYKQYEDMVVESCGR 455  
Db 303 NNVYKQYEDMVVESCGR 321  
  
RESULT 6  
AAW26591  
ID AAW26591 standard; Protein; 321 AA.  
XX  
XX AAW26591;  
AC  
XX  
XX 21-JAN-1998 (first entry)  
DT  
XX Human bone morphogenetic protein BMP-13.  
DE  
XX  
XX BMP-13; bone morphogenetic protein; human; tendon; ligament;  
KW wound healing; tissue repair; tendonitis; human; tendon; ligament;  
KW therapy.  
KW  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FT Peptide 1..201  
FT /label= Sig\_peptide  
FT 202..321  
FT Protein /label= Mat\_protein  
FT /note= "Claim 5"  
FT

FT Protein 218..294  
XX /note= "Claim 5"  
XX US5658882-A.  
XX  
XX 19-AUG-1997.  
XX  
XX 07-DEC-1993; 93US-0164103.  
XX  
XX 22-DEC-1994; 94US-0362670.  
XX 07-DEC-1993; 93US-0164103.  
XX 25-MAR-1994; 94US-0217780.  
XX 02-NOV-1994; 94US-0333576.  
XX (GEM ) GENETICS INST INC.  
XX (HARD ) HARVARD COLLEGE.  
XX Celeste AJ, Melton DA, Rosen VA, Thomsen GH, Wolfman NM;  
PI Wozney JW;  
XX  
XX WPI; 1997-424270/39.  
DR N-PSDB; AAT90386.  
XX  
XX Inducing tendon and ligament formation using BMP-12, BMP-13 or MP-52  
PT - useful for tissue healing and repair, treatment of tendonitis,  
PT Improving fixation of tendons to bone etc  
XX  
XX Claim 5; Column 57-60; 43pp; English.  
XX  
XX This polypeptide comprises a novel bone morphogenetic protein,  
CC designated BMP-13 that induces tendon and ligament formation. Its  
CC amino acid sequence was deduced from isolated genomic clone vl-1  
CC (see AAT90386). A claimed method for inducing formation of tendon  
CC and/or ligament tissues involves the administration of a  
CC composition containing at least one protein selected from BMP-13,  
CC MP52 (see AAW26590) and BMP-12 (see AAW26589). The method is used for  
CC tissue (including skin) healing and repair. This is useful for  
CC treating tendonitis, carpal tunnel syndrome and other defects of  
CC traumatic or congenital origin, in cosmetic surgery and to improve  
CC fixation of tendons or ligaments to bone. The specified proteins  
CC can also be used to increase activity of other BMPs e.g. BMP-2  
CC (see AAW26597).  
XX  
SQ Sequence 321 AA;  
  
Query Match 71.4%; Score 1747; DB 18; Length 321;  
Best Local Similarity 100.0%; Pred. No. 2.7e-140;  
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 137 DLSHTPLRRQKYLFDVSMLSKDELVGAEELRLFRQAPSPAPGPPAGPLHVQLFPCLSPLL 196  
Db 3 DLSHTPLRRQKYLFDVSMLSKDELVGAEELRLFRQAPSPAPGPPAGPLHVQLFPCLSPLL 62  
  
QY 197 LDARTLDPQAGPAGWEVDFVWQGLRHQPKQLCLELRAANGELDAGEAEARAGPQQPP 256  
Db 63 LDARTLDPQAGPAGWEVDFVWQGLRHQPKQLCLELRAANGELDAGEAEARAGPQQPP 122  
  
QY 257 PPDLSLGFGRVRPPQERALLVVFTRSORKNLFAEMREQLGSAEAGPGAGAGSWPPP 316  
Db 123 PPDLSLGFGRVRPPQERALLVVFTRSORKNLFAEMREQLGSAEAGPGAGAGSWPPP 182  
  
QY 317 SGAPDARPLWPSGGRRRRTAFASRHGKRHGKKSRLRCSKKPLHVNFKELGWDWIIAPL 376  
Db 183 SGAPDARPLWPSGGRRRRTAFASRHGKRHGKKSRLRCSKKPLHVNFKELGWDWIIAPL 242  
  
QY 377 EYEAHCEGVCDFFLRSHLEPTNHAIIQTLMNSMDPGSTPPSCCVPTKLTPIISILIYIDAG 436  
Db 243 EYEAHCEGVCDFFLRSHLEPTNHAIIQTLMNSMDPGSTPPSCCVPTKLTPIISILIYIDAG 302  
  
QY 437 NNVYKQYEDMVVESCGR 455  
Db 303 NNVYKQYEDMVVESCGR 321

```

RESULT 7
AAE10982
ID AAE10982 standard; Protein; 321 AA.
XX
AC AAE10982;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human full length VL-1 or BMP-13 protein.
XX
KW Human; bone morphogenic protein; BMP-13; vulnerary; antiinflammatory;
KW analgesic; gene therapy; transforming growth factor-beta; TGF-beta;
KW tissue formation; wound healing; tissue repair; ligament defect; VL-1;
KW carpal tunnel syndrome; tendonitis.
XX
OS Homo sapiens.
XX
FH Key
FT Peptide
FT /label= Signal_peptide
FT Protein
FT /label= Mature_BMP_12_protein
XX
US6284872-B1.
XX
XX 04-SEP-2001.
XX
XX 28-FEB-1997; 97US-0808324.
XX
XX 22-DEC-1994; 94US-0362670.
XX 07-DEC-1993; 93US-0164103.
XX 25-MAR-1994; 94US-0217780.
XX 02-NOV-1994; 94US-0333576.
XX
XX (GEMY ) GENETICS INST INC.
XX (HARD ) HARVARD COLLEGE.
XX
XX Celeste AJ, Wozney JM, Rosen VA, Wolfman NM, Thomsen GH;
XX Melton DA;
XX
XX WPI: 2001-588978/66.
XX N-PSDB; AAD18333.
XX
XX New chimeric DNAs, useful for treating tendonitis, carpal tunnel
XX syndrome and other tendon and ligament defects, comprises DNA encoding
XX propeptide linked to DNA encoding bone morphogenetic proteins (BMP)-12,
XX BMP-13 or MP52 -
XX
XX Example 1; Column 57-60; 42pp; English.
XX
XX The invention relates to a chimeric DNA comprising a DNA sequence
XX encoding a propeptide from a member of the transforming growth factor
XX (TGF)-beta superfamily of proteins. This DNA is linked to a DNA sequence
XX encoding an amino acid sequence encoding a mature polypeptide consisting
XX of Bone Morphogenetic Protein (BMP)-12, BMP-13 or MP52 protein. The DNA
XX sequences are useful for producing proteins which induce tendon/ligament
XX like tissue formation, and for isolating and cloning further DNA
XX sequences encoding BMP-12 related proteins with similar activity. The
XX proteins are useful for the induction of tendon/ligament-like tissue
XX formation, wound healing, ligament and other tissue repair, augmenting
XX the activity of bone morphogenetic proteins, and for treating tendonitis,
XX carpal tunnel syndrome and other tendon and ligament defects. The
XX present sequence is human VL-1 protein also designated as BMP-13.
XX
XX Sequence 321 AA:
XX
XX Query Match
XX Best Local Similarity 100.0%; Pred. No. 2.7e-140;
XX Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 137 DLSHTPLRRQKYLFDVSMLSKDELVGAELRFRQAPSPAGPLHVQLFPCLSPLL 196
XX |

```

```

Db 3 DLSHTPLRRQKYLFDVSMLSKDELVGAELRFRQAPSPAGPLHVQLFPCLSPLL 62
QY 197 LDARTLDPOGAPPAGWEVDVWQGLRHQPKQLCLELRAWGLDAGEAEARAGPOOPP 256
Db 63 LDARTLDPOGAPPAGWEVDVWQGLRHQPKQLCLELRAWGLDAGEAEARAGPOOPP 122
QY 257 PDLRSILGFGRRVRRPQERALLVVFTRSQKKNLFAEMREQLGSAEAGPGAGAGSWPPP 316
Db 123 PDLRSILGFGRRVRRPQERALLVVFTRSQKKNLFAEMREQLGSAEAGPGAGAGSWPPP 182
QY 317 SGAPDARPNLPSGRRRRRTAFASRHKRGKSKKPLHVNFKELGWDWIAPL 376
Db 183 SGAPDARPNLPSGRRRRRTAFASRHKRGKSKKPLHVNFKELGWDWIAPL 242
QY 377 EYEAYHCEGVCDFPLSHLEPTNHAIIOTLMSMDPGSTPPSCCVPTKLTPISTLYIDAG 436
Db 243 EYEAYHCEGVCDFPLSHLEPTNHAIIOTLMSMDPGSTPPSCCVPTKLTPISTLYIDAG 302
QY 437 NNVYKQYEDMVVESGCR 455
Db 303 NNVYKQYEDMVVESGCR 321
XX
XX AAR78739 standard; Protein; 263 AA.
XX
AC AAR78739;
XX
DT 23-NOV-1995 (first entry)
DE Murine mv2 protein.
XX
KW Bone morphogenetic protein; mv2; tendon; ligament.
XX
OS Mus musculus.
XX
FH Key
FT Misc-difference 54
FT /label= P,T
XX
XX WO9516035-A.
XX
XX 15-JUN-1995.
XX
XX 06-DEC-1994; 94WO-US14030.
XX
XX 02-NOV-1994; 94US-0333576.
XX 07-DEC-1993; 93US-0164103.
XX 25-MAR-1994; 94US-0217780.
XX
XX (GEMY ) GENETICS INST INC.
XX (HARD ) HARVARD COLLEGE.
XX
XX Celeste AJ, Melton DA, Rosen VA, Thomsen GH, Wolfman NM;
XX Wozney JM;
XX
XX WPI: 1995-224320/29.
XX N-PSDB; AAQ96224.
XX
XX Bone morphogenetic proteins -12 and -13 and corresp. DNA - used in
XX compsn. for inducing tendon/ligament-like tissue formation
XX
XX Example; Page 71-72; 84pp; English.
XX
XX Oligos #6 and #7 (AAQ96218 & AAQ96219) are used as primers for the
XX amplification of a 275 bp DNA probe, the internal 269 bp of which
XX corresp. to nts #607 to #865 of AAQ96207, from the BMP-12 encoding
XX plasmid subclone PCR1-1#2. This probe was radioactively labelled
XX and used to screen a murine genomic library. DNA sequence analysis
XX of one of positively hybridising recombinants named MVR32 indicates
XX that it encodes a portion of the mouse gene corresp. to the PCR
XX product mv2 (murine homolog of the human VL-1 sequence AAQ96213.

```

XX  
BB  
CC-DEC-1004-  
0490-0300670

DR N-P;

DR N-P;

```
PR 07-DEC-1993; 93US-0164103.
PR 25-MAR-1994; 94US-0217780.
PR 02-NOV-1994; 94US-0333576.
XX (GENY ) GENETICS INST INC.
PA (HARD ) HARVARD COLLEGE.
XX Celeste AJ, Wozney JM, Rosen VA, Wolfman NM, Thomsen GH;
PI Melton DA;
XX WPI; 2001-588978/66.
DR N-PSDB; AAD18336.
XX New chimeric DNAs, useful for treating tendonitis, carpal tunnel
PT syndrome and other tendon and ligament defects, comprises DNA encoding
PT propeptide linked to DNA encoding bone morphogenetic proteins (BMP)-12,
PT BMP-13 or MP52 -
XX Example 1; Column 71-74; 42pp; English.
PS The invention relates to a chimeric DNA comprising a DNA sequence
XX encoding a propeptide from a member of the transforming growth factor
CC (TGF)-beta superfamily of proteins. This DNA is linked to a DNA sequence
CC encoding an amino acid sequence encoding a mature polypeptide consisting
CC of Bone Morphogenetic Protein (BMP)-12, BMP-13 or MP52 protein. The DNA
CC sequences are useful for producing proteins which induce tendon/ligament
CC like tissue formation, and for isolating and cloning further DNA
CC sequences encoding BMP-12 related proteins with similar activity. The
CC proteins are useful for the induction of tendon/ligament-like tissue
CC formation, wound healing, ligament and other tissue repair, augmenting
CC the activity of bone morphogenetic proteins, and for treating tendonitis,
CC carpal tunnel syndrome and other tendon and ligament defects. The
CC present sequence is murine partial mv2 protein which is homologous to
CC human BMP-12 or VL-1 sequences of the invention.
XX Sequence 263 AA;
SQ
Query Match 49.1%; Score 1202.5; DB 22; Length 263;
Best Local Similarity 83.3%; Pred. No. 4.3e-94;
Matches 230; Conservative 6; Mismatches 21; Indels 19; Gaps 4;
QY 180 PAGPLHVQLFPCLLDARTLDPOGAPPAGWVEFDVWQGLRHQPWQLCLELRAANGE 239
DB 7 PAG-----PTLRGSSQTQR-----PAG-KSFVDVWQGLRHQPWQLCLELRAANGE 51
QY 240 LDAGEAARAGQPQPPDLRLSGRGRVRPPQERALLVFTSRQKNLFAEMREOLGS 299
DB 52 LDXTGARAGQPQPPDLRLSLGFRGRVRPPQERALLVFTSRQKNLFTMEHEQLGS 111
QY 300 ABAAGPGAGAGSWPPPGAPDARPLPSGRRRRRTAFASRHKRHKKSRLCRSKPL 359
DB 112 AEA----AGAEGSCAPSGSDTGSWLPSPGRRRRRTAFASRHKRHKKSRLCRSKPL 167
QY 360 HVNFKELGDDWIIAPLEYAHCVCVDFPLRSHLEPTNHAIITLNMNSMDPGSTPPSC 419
DB 168 HVNFKELGDDWIIAPLEYAHCVCVDFPLRSHLEPTNHAIITLNMNSMDPGSTPPSC 227
QY 420 CVPTKLTPTISILYIDAGNNVYKQYEDMVVESCGR 455
DB 228 CVPTKLTPTISILYIDAGNNVYKQYEDMVVESCGR 263
RESULT 11
ID AAR69600 standard; Protein; 501 AA.
XX AAR69600;
AC AAR69600;
XX 10-OCT-1995 (first entry)
DT New TGF-beta family member - MP-52 protein sequence.
DE Transforming growth factor-beta family; mitogenic; differentiation;
XX
```

```
KW treatment; prevention; disease; bone; cartilage; connective tissue;
KW skin; mucosa; epithelium; dental tissue; wound healing; osteoporosis;
XX tissue regeneration; arthritis; ss.
OS Homo sapiens.
XX Key Location/Qualifiers
FH Peptide 382
FT /label= mature protein
XX
XX WO9504819-A.
XX 16-FEB-1995.
XX 09-AUG-1994; 94WO-EP02630.
XX 10-AUG-1993; 93DE-4326829.
PR 25-MAY-1994; 94DE-4418222.
PR 09-JUN-1994; 94DE-4420157.
XX (BIOP-) BIOPHARM GES BIOTECHNOLOGISCHEN ENTWICKL.
XX Hotten G, Neidhardt H, Paulista M, Hoetten G;
XX WPI; 1995-090897/12.
XX N-PSDB; AAQ83695.
XX New DNA encoding a new member of the TGF beta family - and
PT related vectors, host cells etc., has mitogenic and
PT differentiation inducing activity, e.g. for treating or
PT preventing diseases of bone and cartilage etc.
XX Claim 6; Page 36; 51pp; German.
XX The amino acid sequence of a novel member of the transforming growth
CC factor-beta (TGF-b) family named MP-52. The gene encodes a protein
CC of 501 amino acids (AA). The protein, or at least the mature protein,
CC has mitogenic and/or differentiation inducing properties useful in
CC the treatment or prevention of diseases of bone, cartilage, connective
CC tissue, skin, mucosa, epithelium or dental tissue. The protein can also
CC be used for wound healing and tissue regeneration e.g. in osteoporosis
CC and arthritis.
XX Sequence 501 AA;
SQ
Query Match 37.9%; Score 927.5; DB 16; Length 501;
Best Local Similarity 44.7%; Pred. No. 2.3e-70;
Matches 209; Conservative 64; Mismatches 98; Indels 97; Gaps 16;
QY 38 GSTKGMRSRKEGKMQRAPRDSADGREGQEPQPRQDEPRAQO----- 79
DB 81 GQTGGLTQPKDEPKLP-----PRPG-GPEPKPGHPQTRQATARTVTPKQLPGGKAP 134
QY 80 -----PRAQEPGRGRVPRVPEHYMLSTYRTSYSTAELGIGNASFFQ 119
DB 135 PRAGSVSPSSFLKKAREPGPPREPKEPRPPPTTPEHYMLSLYRTLSADRGKGNSSVKL 194
QY 120 SSKSANTITSFVDRGLDLSHTPLRROKYLFDVSMLSKEELVGAELRFRQAPS----- 174
DB 195 EAGLANTITSFIDKGDGRGPV-VRKQRTYFDISAL-EKDGLLGAELRLRLKRPSTAKP 252
QY 175 -APWGPAGPLHVQLFPCLS-----PLLLDARTLDPOGAPPAGWVEFDVWQGLRH-QPWKQ 228
DB 253 AAPGGGRAAQL--KLSSCPSGRQPSALLDVRSV--PGLDGGSGWEFIDKWLFRFNKNSAQ 308
QY 229 LCLELRAAWGELDAGEAEARAGQPQPPDLRLSGFRGRVRPPQERALLVFTSRQKN 288
DB 309 LCLELE-AW---ERGRA-----VDLURLGLFDRAARQVHEKALFLVFGRTKKRD 352
QY 289 L-FAEMREOLGSAAEAGPGAGAGSWPPPGAPDARPLPSGRRRRRTAFASRHKRHK 347
DB 353 LFFNEIKARSGQDDKT-----VVEYLFQ-RRRRRAPLATROGKRPS 393
```



```
QY 348 KKSRLRCSKKPLHVNFKELGWDWIIAPLEYEAYHCEGVCDPFLRSHLEPTNHAIQTLM 407
| : |||:| |||||:|||||||:||||:||||:|||||||:|||||
Db 394 KNLKARCSKALHVNFKDMGWDWIIAPLEYEAFHCEGLCEPFLRSHLEPTNHAIQTLM 453
QY 408 NSMDPGSTPPSCVPTKLTPIISILYIDAGNNVVYKQYEDMVVESCGR 455
||||| |||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 454 NSMDPESTPTCCVPTRLSPISILFIDSANNVVYKQYEDMVVESCGR 501

RESULT 12
AAW36100
ID AAW36100 standard; Protein; 501 AA.
XX
AC AAW36100;
XX
DT 08-MAY-1998 (first entry)
XX
DE Human MP52.
XX
KW Bone morphogenetic protein; BMP; processing enzyme; MP52;
KW BMP-2; BMP-4; BMP-6; BMP-7; bone formation; bone regeneration.
XX
OS Homo sapiens.
XX
PN WO9741250-A1.
XX
PD 06-NOV-1997.
XX
PF 28-APR-1997; 97WO-JP01474.
XX
PR 30-APR-1996; 96JP-0130618.
XX
PA (FARH ) HOECHST YAKUHIN KOGYO KK.
PA (FARH ) HOECHST PHARM & CHEM KK.
XX
PI Kimura M, Makishima F, Takahashi M;
XX
WPI; 1997-549748/50.
DR N-PSDB; AAT98191.
XX
PT Production of mature bone morphogenetic protein - by treatment of
PT precursor protein with a processing enzyme such as furin either
PT directly or by expressing them both in the same host
XX
PS Example 1; Pages 21-25; 34pp; Japanese.
XX
CC The present sequence is MP52, which is a bone morphogenetic
CC protein (BMP).
CC Mature BMP can be produced by directly adding a BMP processing
CC enzyme to a solution containing BMP precursor protein, or by
CC transforming an animal cell with expression vectors containing DNA
CC encoding the enzyme and precursor protein, culturing the
CC transformant and isolating the mature BMP from the culture. The
CC method can be used to produce MP52, BMP-2, BMP-4, BMP-6 and BMP-7,
CC which can be used to treat bone formation or regeneration
CC abnormalities.
XX
SQ Sequence 501 AA;

Query Match 37.9%; Score 927.5; DB 18; Length 501;
Best Local Similarity 44.7%; Pred. NO. 2.3e-70;
Matches 209; Conservative 64; Mismatches 98; Indels 97; Gaps 16;

QY 38 GSTKGRSRKCKMORAPRDSAGREGQPPQPRQDEBRAQ----- 79
| | | | | : : : : | | | | | : : | : | : | : | : | : |
Db 81 GGTGLTPQPKDEPKL-----PRPG-GPEPKPGHPQTRQATARTVTPKQLPGGRAP 134
QY 80 -----PRAQPPGPRWPPEHYMLSIYRTYSTAELKGINASFPQ 119
| | | | | : | | | | | : |||||:|||||:|||||:|||||
Db 135 PKAGSVSPSFLKKAREPCPPREPFPPTTHEYMLSIYRLSDADRGKGNSSVKL 194
QY 120 SKSANTITSFVDRGLDLSHTPLRRQRKYLFDVMSLSEELVGAEIRLFRQAPS----- 174
: |||||:|:| | | :|:|:|:|:| | :|:|:|:|:| | | | |
```

```
Db 195 EAGLANTIITSFIDKGODDRGPV-VRKQRYVFDISAL-EKDGLLGAEILRLRKKPSDTAKP 252
QY 175 -APWGPAGPLHVLQFPCLs-----PLLLDARTLDPOGAPPAGWEVDVWQGLRH-QPWKQ 228
| | | | | : | | | | | : |||:|:| | | : | | | | | : | : |
Db 253 APGGGAAQL--KLSSCFSGROPASLLDVRSV--PGLDGSWEVDIWKLFNFKNSAQ 308
QY 229 LCLELRAAWGELDAGEAEARAGPOQPPPLRLSLGFGRRVRPPQBRALLVVFTRSORKN 288
||||| || : | | | | | : ||| ||| | | | | | : |||:|:|:|:|:|
Db 309 LCLELE-AW---ERGRA-----VDLUGLGFDRAAQVHEKALFLVFGRTKKRD 352
QY 289 L-FAEMREQLGSAEAAAGPCAGAGSGWPPSPGAPDARWLPSPGRRRRRTAFASRHGKRHG 347
| | | : | : | : | : | : | : | : | : | : | : | : | : |
Db 353 LFFNEIKARSGQDDKT-----VVEYLFQ-RRKRRAPLATRQGRKRS 393
QY 348 KKSRLRCSKKPLHVNFKELGWDWIIAPLEYEAYHCEGVCDPFLRSHLEPTNHAIQTLM 407
| : |||:| |||||:|||||||:||||:||||:|||||||:|||||
Db 394 KNLKARCSKALHVNFKDMGWDWIIAPLEYEAFHCEGLCEPFLRSHLEPTNHAIQTLM 453
QY 408 NSMDPGSTPPSCVPTKLTPIISILYIDAGNNVVYKQYEDMVVESCGR 455
||||| |||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 454 NSMDPESTPTCCVPTRLSPISILFIDSANNVVYKQYEDMVVESCGR 501

RESULT 13
AAW19210
ID AAW19210 standard; Protein; 501 AA.
XX
AC AAW19210;
XX
DT 04-MAR-1998 (first entry)
XX
DE Human TGF-beta protein MP52.
XX
KW Human transforming growth factor-beta; TGF-beta; MP52; superfamily;
KW cartilage; bone inducing activity; inhibit; bone resorption.
XX
OS Homo sapiens.
XX
PN DE19548476-A1.
XX
PD 26-JUN-1997.
XX
PF 22-DEC-1995; 95DE-1048476.
XX
PR 22-DEC-1995; 95DE-1048476.
XX
PA (BIOP-) BIOPHARM GES BIOTECHNOLOGISCHEN ENTWICKL.
XX
PI Bechtold R, Hotten G, Paulista M, Pohl J, Hoetten G;
XX
WPI; 1997-333931/31.
DR N-PSDB; AAT69695.
XX
PT Compound containing protein from TGF-beta superfamily - has bone
PT and/or cartilage inducing activity, useful in treatment of, e.g.
PT osteoporosis, bone damage, Paget's disease and osteoarthritis
XX
PS Claim 3; Page 9; 10pp; German.
XX
CC This sequence is the human transforming growth factor (TGF)-beta protein
CC designated MP52. MP52 can be used in a compound of formula (I):
CC A-X(1-20)-B(1-20) (I); A = protein, or fragment, of the TGF-beta
CC superfamily with cartilage and/or bone inducing activity (e.g. MP52);
CC B = 1 or more substituent groups with an affinity to the extracellular
CC matrix, cellular components of bone and/or cartilage and/or to a
CC biocompatible carrier matrix; X = 1 or more covalent bonds and/or spacer
CC groups. The compound may be used to inhibit bone resorption, prevent or
CC treat bone or cartilage related disorders, including osteoporosis,
CC Paget's disease, osteodystrophy, osteoarthritis or osteoarthritis and
CC to treat bone or cartilage damage caused by wounding or overloading.
XX
SQ Sequence 501 AA;
```

Query Match 37.9%; Score 927.5; DB 18; Length 501;  
Best Local Similarity 44.7%; Pred. No. 2.3e-70;  
Matches 209; Conservative 64; Mismatches 98; Indels 97; Gaps 16;

QY 38 GSTKMRKKEKQAPRDSAGREGQEPQRPDEPRAQO----- 79  
| | : : : : | | : : : : | | : : : : | | : : : :  
DB 81 GOTGLTPKPKDEPKLP-----PRPG-GPEPKGHPQTRQATARTVTPKQLPGGKAP 134  
QY 80 -----PRAQPPGGRPRVPHVEYMLSYRTSYAEKLGINASFQ 119  
| | : : : : | | : : : : | | : : : : | | : : : :  
DB 135 PKAGSVPSFLLKKAREPGPREPEPRPPITPHEYMLSLYRTLSADRRKGGNSVKL 194  
QY 120 SSKSANTITSFVDRGLDLSHTPLRRQKYLFDVSMLSKEELVGAEELRFOAPS----- 174  
| | : : : : | | : : : : | | : : : : | | : : : :  
DB 195 EAGLANTITSFIDKQDGRGPV-VKQRYVFDISAL-EKDGLLGAELRLRKKPSDTAKP 252  
QY 175 -APWGPPAGPLHVLPCLSL-----PLLLDARTLDPOGAPPAGWEFVWQGLRH-OPWKQ 228  
| | : : : : | | : : : : | | : : : : | | : : : :  
DB 253 AAPGGGAAQL--KLSSCPSGRQPASLLDVRSV--PGLDGGSGWEFVWKLFRNFKNSAQ 308  
QY 229 LCLELRAANGELDAGEAEARARGPOPPDLRLSLGFGRRVPPPOERALLVVFTSRQKN 288  
| | : : : : | | : : : : | | : : : : | | : : : :  
DB 309 LCLELE-AW---ERGRA-----VDLRGLGFDRARQVHEKALFLVFGTKKRD 352  
QY 289 L-FAEMREQLGSAEAGPCAGAGSWPPSPGAPDARPLPSGRRRRRTAFASRHKRHH 347  
| | : : : : | | : : : : | | : : : : | | : : : :  
DB 353 LFFNEIKARSGDDKT-----VVEYLFQO-RRKRAPLATRQGRKPS 393  
QY 348 KKSRLRCSKKPLHVNFKELGWDDWIIAPLEYAYHCEGVCDFFPLRSHLEPTNHAIQTLM 407  
DB 394 KNLKARCSRKALHVNFKDMGWDDWIIAPLEYEAFHCEGLCEPPLRSHLEPTNHAVIQTLM 453  
QY 408 NSMDPGSTPPSCCVPTKLTPIISILYIDAGNNVYKQYEDMVVESCGR 455  
| | : : : : | | : : : : | | : : : : | | : : : :  
DB 454 NSMDPESTPPTCCVPTRLSPISILFIDSANNVYKQYEDMVVESCGR 501

RESULT 15  
AAW01799  
ID AAW01799 standard; Protein; 501 AA.  
XX  
AC AAW01799;  
XX  
DT 15-OCT-1997 (first entry)  
XX  
DE Human MP52 protein.  
XX  
KW Human; MP52; transforming growth factor; TGF; beta; medicament;  
KW treatment; prevention; nervous system; disease; neuropathology;  
KW ageing.  
XX  
OS Homo sapiens.  
XX  
PN DEI9525416-A1.  
XX  
PD 16-JAN-1997.  
XX  
PF 12-JUL-1995; 95DE-1025416.  
XX  
PR 12-JUL-1995; 95DE-1025416..

Human high mol. wt. protein MP52, a growth/differentiation factor.  
Growth factor; differentiation; bone induction; osteoporosis; teeth;  
tooth; dental; joint tissue; cartilage; mucous membrane; skin; nails;  
wound healing; regeneration; skeletal disorder; fracture; dimer.  
Homo sapiens.  
WO9704095-A1.  
06-FEB-1997.  
24-JUL-1996; 96WO-JP02065.  
24-JUL-1995; 95JP-0218022.  
(FARH ) HOECHST JAPAN LTD.  
(FARH ) HOECHST PHARM & CHEM KK.  
Fujino Y, Kawai S, Kimura M, Matsumoto T, Takahashi M;  
WPI; 1997-132636/12.  
N-ESDB; AAT61412.  
High molecular weight human MP52 growth or differentiation factor -  
promotes bone induction, is useful for treatment and prevention of  
bone disease

PS Claim 1; Page 12-16; 25pp; Japanese.  
XX  
CC AAW11900 is a high mol. wt. form of a human growth/differentiation  
CC factor MP52. MP52 promotes bone induction and is useful for plastic  
CC reconstructive surgery, cosmetic facial treatment, bone transplantation  
CC and tooth implantation. It is also useful for the treatment and  
CC prevention of disorders of bone formation, bone, cartilage, joint  
CC tissue, skin, mucous membranes, nails or teeth; for wound treatment and  
CC tissue regeneration; and for the treatment of skeletal disorders and  
CC fractures.  
XX  
SQ Sequence 501 AA;  
Query Match 37.9%; Score 927.5; DB 18; Length 501;  
Best Local Similarity 44.7%; Pred. No. 2.3e-70;  
Matches 209; Conservative 64; Mismatches 98; Indels 97; Gaps 16;

QY 38 GSTKMRKKEKQAPRDSAGREGQEPQRPDEPRAQO----- 79  
| | : : : : | | : : : : | | : : : : | | : : : :  
DB 81 GOTGLTPKPKDEPKLP-----PRPG-GPEPKGHPQTRQATARTVTPKQLPGGKAP 134  
QY 80 -----PRAQPPGGRPRVPHVEYMLSYRTSYAEKLGINASFQ 119  
| | : : : : | | : : : : | | : : : : | | : : : :  
DB 135 PKAGSVPSFLLKKAREPGPREPEPRPPITPHEYMLSLYRTLSADRRKGGNSVKL 194  
QY 120 SSKSANTITSFVDRGLDLSHTPLRRQKYLFDVSMLSKEELVGAEELRFOAPS----- 174  
| | : : : : | | : : : : | | : : : : | | : : : :  
DB 195 EAGLANTITSFIDKQDGRGPV-VKQRYVFDISAL-EKDGLLGAELRLRKKPSDTAKP 252  
QY 175 -APWGPPAGPLHVLPCLSL-----PLLLDARTLDPOGAPPAGWEFVWQGLRH-OPWKQ 228  
| | : : : : | | : : : : | | : : : : | | : : : :  
DB 253 AAPGGGAAQL--KLSSCPSGRQPASLLDVRSV--PGLDGGSGWEFVWKLFRNFKNSAQ 308  
QY 229 LCLELRAANGELDAGEAEARARGPOPPDLRLSLGFGRRVPPPOERALLVVFTSRQKN 288  
| | : : : : | | : : : : | | : : : : | | : : : :  
DB 309 LCLELE-AW---ERGRA-----VDLRGLGFDRARQVHEKALFLVFGTKKRD 352  
QY 289 L-FAEMREQLGSAEAGPCAGAGSWPPSPGAPDARPLPSGRRRRRTAFASRHKRHH 347  
| | : : : : | | : : : : | | : : : : | | : : : :  
DB 353 LFFNEIKARSGDDKT-----VVEYLFQO-RRKRAPLATRQGRKPS 393  
QY 348 KKSRLRCSKKPLHVNFKELGWDDWIIAPLEYAYHCEGVCDFFPLRSHLEPTNHAIQTLM 407  
DB 394 KNLKARCSRKALHVNFKDMGWDDWIIAPLEYEAFHCEGLCEPPLRSHLEPTNHAVIQTLM 453  
QY 408 NSMDPGSTPPSCCVPTKLTPIISILYIDAGNNVYKQYEDMVVESCGR 455  
| | : : : : | | : : : : | | : : : : | | : : : :  
DB 454 NSMDPESTPPTCCVPTRLSPISILFIDSANNVYKQYEDMVVESCGR 501

RESULT 15  
AAW01799  
ID AAW01799 standard; Protein; 501 AA.  
XX  
AC AAW01799;  
XX  
DT 15-OCT-1997 (first entry)  
XX  
DE Human MP52 protein.  
XX  
KW Human; MP52; transforming growth factor; TGF; beta; medicament;  
KW treatment; prevention; nervous system; disease; neuropathology;  
KW ageing.  
XX  
OS Homo sapiens.  
XX  
PN DEI9525416-A1.  
XX  
PD 16-JAN-1997.  
XX  
PF 12-JUL-1995; 95DE-1025416.  
XX  
PR 12-JUL-1995; 95DE-1025416..

```
XX (BIOP-) BIOPHARM GES BIOTECHNOLOGISCHEN ENTWICKL.
XX
XX PI Bechtold R, Hoetten G, Paulista M, Pohl J, Unsicker K;
XX
XX DR WPI: 1997-078343/08.
XX DR N-PSDB; AAT59405.
XX
XX PT Medicaments contg. protein MP52 - useful for treating neurological
XX PT disorders
XX
XX PS Claim 2; Pages 12-14; 21pp; German.
XX
XX CC The present sequence is the human MP52 protein, which is
XX CC described in WO 9316099 and 9504819 as a member of the human
XX CC transforming growth factor beta superfamily. Active MP52 can be
XX CC used in a medicament to treat and prevent nervous system diseases,
XX CC and/or to treat neuropathological conditions caused by nervous
XX CC system ageing.
XX
XX SQ Sequence 501 AA;
Query Match 37.9%; Score 927.5; DB 18; Length 501;
Best Local Similarity 44.7%; Pred. No. 2.3e-70;
Matches 209; Conservative 64; Mismatches 98; Indels 97; Gaps 16;
QY 38 GSTKGMRSRKEGKMQRAPRDSAGREGQEPQPRPQDEPRAQQ----- 79
DQ 81 GQTGGLTPQPKDEPKLP-----PRPG-GPEPKPGHPQTQATARTVTPKGOLPGGKAP 134
QY 80 -----PRAQEPGPGPRVVPHEYMLSYRTYSIAEKLGINASFTQ 119
DQ 135 PKAGSVPSFLKKAREPGPREPKPEPRPPITPHEYMLSLYRTLSDADRKGNSSVKL 194
QY 120 SSKSANTITSFVDRGLDLSHTPLRRQKYLEVDVMSLSDKEELVGAELRLFRQAPS----- 174
DQ 195 EAGLAWTITSFIDKGODDRGPV--VRKQRYVFDISAL-EKDGLLGAELRLIRKKPSDTAKP 252
QY 175 -APWGPAGPLHVQLFPCLSL---PLLLDARTLDQAGPAGWEFVDMQGLRH-OPWKQ 228
DQ 253 AAPGGGAAQL--KLSSCFSGRQPASLLDVRSV--PGLDGSWEVFDIWKLPFNKNSAQ 308
QY 229 LCLELRAANGELDAGEAARAGPQQPPDLRLSLGFGRRVRPPQERALLVVTSRQKN 288
DQ 309 LCLELE-AW---ERGRA-----VDLRGLGFDRAARQVHEKALFLVFGRTKKRD 352
QY 289 L-FAEMREOLGSAEAGPGAGAGSWPPSPGAPDARPWLSPGRRRRRTAFASRHGKRHG 347
DQ 353 LFFNEIKRSGQDDKT-----VYEIFSQ-RRRRAPLATROGKRPS 393
QY 348 KKSRLRCRKKPLHVNFKELGWDWIIAPLEYEAYHCEGVCDFPLRSHLEPTNHAIQTLM 407
DQ 394 KNLKARCSRKALHVNFKDNGWDWIIAPLEYEAFHCEGLCEFPRLSHLEPTNHAVIQTLM 453
QY 408 NSMDPGSTPPSCCVPKLTPISILYIDAGNNVYKQYEDMVVESCGR 455
DQ 454 NSMDPESTPTCCVPRLSPIILFIDSANNVYKQYEDMVVESCGR 501
```

Search completed: November 25, 2002, 02:55:19  
Job time : 84 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 00:59:55 ; Search time 2432 Seconds

(without alignments)

9489.543 Million cell updates/sec

Title: US-09-825-751A-19

Perfect score: 1425

Sequence: 1 ctctcgggagacgagcaca.....tttccgcgcgttgccgcg 1425

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: gb\_gss:\*

18: em\_gss\_hum:\*

19: em\_gss\_inv:\*

20: em\_gss\_pln:\*

21: em\_gss\_vrt:\*

22: em\_gss\_fun:\*

23: em\_gss\_nam:\*

24: em\_gss\_mus:\*

25: em\_gss\_other:\*

26: em\_gss\_pro:\*

27: em\_gss\_rod:\*

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	534.2	37.5	543	13	BI559439
2	345	24.2	450	10	BE098696
3	303.4	21.3	403	9	AI535156
4	271.4	19.0	563	12	BR041119
5	253.6	17.8	661	14	BQ169372
6	250.2	17.6	343	12	BF542757

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

RESULT 1  
BI559439  
LOCUS 603252937F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:5295441 5',  
DEFINITION mRNA sequence.  
ACCESSION BI559439  
VERSION BI559439.1 GI:15446753  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 543)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
DNA Sequencing by: Inceye Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM11747 row: 1 column: 10  
High quality sequence stop: 532.  
Location/Qualifiers

ALIGNMENTS

7	237.6	16.7	798	14	BQ604531
8	237	16.6	475	9	AI761284
9	236	16.6	505	13	BI401488
10	230.8	16.2	1025	12	BF588622
11	217.4	15.3	552	9	AI641304
12	209.4	14.7	307	12	BF420157
13	191.8	13.5	479	14	BQ563481
14	191.6	13.4	283	17	AQ311900
15	188.4	13.2	713	10	BE114678
16	186.2	13.1	272	10	BE637633
17	171.4	12.0	597	12	BG162724
18	159	11.2	583	13	BI473851
19	152.4	10.7	608	17	CNS01SYJ
20	144.8	10.2	200	9	AU241555
21	144	10.1	822	13	BI903718
22	144	10.1	841	13	BG69662
23	141.6	9.9	714	12	BF689009
24	139.4	9.8	972	17	CNS0442A
25	138.4	9.7	284	13	BI901630
26	138.4	9.7	316	13	BI901881
27	138.4	9.7	793	12	BG187538
28	138.4	9.7	991	12	BE740547
29	138.2	9.7	525	14	BM691373
30	137.2	9.6	577	12	BF045110
31	136.8	9.6	805	12	BG214178
32	132	9.3	1002	17	CNS04IFG
33	131.4	9.2	656	13	BI400982
34	131	9.2	364	14	BQ567172
35	131	9.2	831	14	BQ715851
36	130.6	9.2	3392	11	BC023526
37	130.2	9.1	563	9	AL788844
38	130.2	9.1	672	14	BQ525130
39	129.6	9.1	855	12	BG186221
40	129.6	9.1	883	9	ALS81402
41	129.6	9.1	968	9	AL572764
42	129.2	9.1	817	9	AL576612
43	129	9.1	875	12	BE893005
44	128.4	9.0	441	12	BG385196
45	128.4	9.0	984	9	AI478076

BQ604531 MI-P-CPI-  
AI761284 wh98504.x  
BI401488 MI-P-CPD-  
BF588622 602185258  
AI641304 fci3c07.y  
BF420157 UI-R-BJ2-  
BQ563481 gi03h02.y  
AQ311900 RPC111-94  
BE114678 UI-R-BJ1-  
BE637633 BB637633  
BG162724 dfl1g04.y  
BI473851 fp47f01.y  
AI165844 Tetraodon  
AU241555 AU241555  
BI903718 603167294  
BG69662 602837409  
BF689009 602185258  
AL273547 Tetraodon  
BI901630 id14g06.x  
BI901881 id14g06.y  
BG187538 RST6663.A  
BE740547 601595643  
BM691373 UI-E-C11-  
BF045110 BP250018A  
BG214178 RST33812  
AL292165 Tetraodon  
BI400982 MI-P-AV1-  
BQ567172 gi80c12.y  
BQ715851 AGENCOURT  
BC023526 Homo sapi  
AL788844 AL788844  
BQ525130 NISC.d009  
BG186221 RST5070.A  
ALS81402 ALS81402  
AL572764 AL572764  
AL576612 AL576612  
BE893005 601434183  
BG385196 306833.MA  
AI478076 MPMGP621\_



```
QY 814 AGCTGGCTTCGGCGGAGGGTGGCGCC 842
|||||
Db 30 AGCTGGCTTCGGCGGAGGGTGGCGCC 2

RESULT 3
AI535156/c
LOCUS
DEFINITION
  UI-R-C3-sq-b-01-0-UI-s1 UI-R-C3 Rattus norvegicus cDNA clone
  UI-R-C3-sq-b-01-0-UI 3', mRNA sequence.
ACCESSION
  AI535156
VERSION
  AI535156.1 GI:4449291
KEYWORDS
  EST.
SOURCE
  Norway rat.
ORGANISM
  Rattus norvegicus
  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
  Rattus.
REFERENCE
  1 (bases 1 to 403)
  Bonaldo,M.F., Lennon,G. and Soares,M.B.
  Normalization and subtraction: two approaches to facilitate gene
  discovery
  Genome Res. 6 (9), 791-806 (1996)
JOURNAL
  9704477
MEDLINE
  Contact: Soares, MB
  Program for Rat Gene Discovery and Mapping
  University of Iowa
  451 Eckstein Medical Research Building Iowa City, IA 52242, USA
  Tel: 319 335 8250
  Fax: 319 335 9565
  Email: msoares@blue.weeg.uiowa.edu
  Oligo-dt track not found, Not 1 site shown in beginning of sequence
  is likely internal to the message. cDNA Library Preparation: M.B.
  Soares Lab Clone distribution: clones will be available through
  Research Genetics (www.resgen.com) This clone is also available
  through the I.M.A.G.E. Consortium at LNL (info@image.lnl.gov).
  IMAGE ID=1768988
Seq primer: M13 Forward
POLYA=No.
Location/Qualifiers
  1..403
  /organism="Rattus norvegicus"
  /strain="Sprague-Dawley"
  /db_xref="taxon:10116"
  /clone="UI-R-C3-sq-b-01-0-UI"
  /clone_lib="UI-R-C3"
  /dev_stage="adult"
  /lab_host="DH10B (Life Technologies)"
  /note="Vector: pT73D-Pac (Pharmacia) with a modified
  polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-C3
  library is a subtracted library of a series, ultimately
  derived from a mixture of individually tagged normalized
  libraries from rat placenta, adult lung, brain, liver,
  kidney, heart, spleen, ovary, muscle, and 8, 12 and 18-day
  embryos, after a series of subtractions to reduce the
  representation of cDNAs from which ESTs had already been
  generated. The following serially subtracted libraries
  were generated in this process: UI-R-C3, UI-R-C2p, UI-R-C1
  , UI-R-C0, UI-R-A1, UI-R-E1. The tag is a string of 3-5
  nucleotides present between the Not I site and the
  oligo-dt track which allows identification of the library
  of origin of a clone within the mixture. The subtracted
  library (UI-R-C3) was constructed as follows: PCR amplified
  cDNA inserts from UI-R-C2p clones from which 3' ESTs had
  been derived was used as a driver in a hybridization with
  the UI-R-C2p library in the form of single-stranded
  circles. The remaining single-stranded circles (subtracted
  library) was purified by hydroxyapatite column
  chromatography, converted to double-stranded circles and
  electroporated into DH10B bacteria (Life Technologies) to
  generate the UI-R-C3 library. This procedure has been
  previously described (Bonaldo, Lennon and Soares, Genome
  Research 6:791-806, 1996)"

BASE COUNT      81 a      130 c      123 g      68 t      1 others
ORIGIN
  Query Match      21.3%; Score 303.4; DB 9; Length 403;
  Best Local Similarity 84.6%; Pred. NO. 1.3e-49;
  Matches 340; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 441 TCCTCGCACACCTCTCCGAGAGACAGAAAGTATTTGTTGATGTGTCATGCTCTCAGA 500
|||||
Db 403 TCCTCGCACACCTCTCCGAGAGACAGAAAGTATTTGTTGATGTGTCATGCTCTCAGA 344
|||||
QY 501 CAAAGAAGAGCTGGTGGCGCGAGCTCGGCTCTTTCGCCAGCGCCCTCAGCGCCTG 560
|||||
Db 343 CAAAGAAGAGCTGGTGGCGCGAGCTCGGCTCTTTCGCCAGCGCCCTCAGCGCCTG 284
|||||
QY 561 GGGGCCACACCGCGCGCGCTCCACGTGCAGCTTTCCTTGCCTTTCGCCCTACTGCT 620
|||||
Db 283 GGGGCCACACCGCGCGCTCCACGTGCAGCTTTCCTTGCCTTTCGCCCTACTGCT 224
|||||
QY 621 GGACGGCGGACCTGTGACCGCGCGCGCGCGCGCGCGCGCTGGGAAGTCTTTCGACGT 680
|||||
Db 223 GGACTTAGGACCTGTGATCTCAGGGGCCAACAGAAGCTGGCTGGGAAGTCTTTCGACGT 164
|||||
QY 681 GTGGCAGGCTCGGCCACCGCCCTGGAAAGCAGCTGTGCTTGAGCTGCGGGCCGCGATG 740
|||||
Db 163 GTGGCAGGCTCGGCCCTCAGCCGTGGAAGCAGCTGTGCTTGAGCTTTCGGCAGTCTG 104
|||||
QY 741 GGGCGAGCTGAGCGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 800
|||||
Db 103 GGGTGAGCTGGACGCCAGGAGTTTCAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCC 44
|||||
QY 801 CCCGGACCTCGGAGTCTGGGCTTCGGCGGAGGCTGGCGGCC 842
|||||
Db 43 CTTGGACCTCGGAGCTGGGCTTCGGCGGAGGCTGGCGGCC 2

RESULT 4
BF044119/c
LOCUS
DEFINITION
  BP250023A10F2 Soares normalized bovine placenta Bos taurus cDNA
  clone BP250023A10F2 5', mRNA sequence.
ACCESSION
  BF044119
VERSION
  BF044119.1 GI:10761174
KEYWORDS
  EST.
SOURCE
  COW.
ORGANISM
  Bos taurus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
  Bovidae; Bovinae; Bos.
  1 (bases 1 to 563)
  Lewin,H.A., Soares,M.B., Rebeiz,M., Pardinas,J., Liu,L. and Larson
  J.H.
  Bovine ESTs
  Unpublished (2000)
  Contact: Lewin, H. A.
  W. M. Keck Center for Comparative and Functional Genomics
  University of Illinois at Urbana-Champaign
  340 Edward R. Madigan Laboratory, 1201 W. Gregory Dr., Urbana, IL
  61801, USA
  Tel: 217 333 5998
  Fax: 217 244 5617
  Email: h-lewin@uiuc.edu
  Funding for cattle EST sequencing was provided by the USDA National
  Research Initiative, Animal Genome Resource Grant AC 99-3205-8534
  to H. A. Lewin and J. E. Womack. Base Calling/Quality Scores: PHRED
  from Washington University Genome Center. Vector Trimmi g:
  Cross-match from Washington University Genome Center PHRAP suite.
  This sequence is vector free and at least 200 bp in length.
  PCR Primers
  FORWARD: TAATACGACTCTACTATAGG
  BACKWARD: ATTAACCTCTACTAAG
  Insert Length: 563 Std Error: 0.00
  Plate: BP250023A10 row: F column: 2
```

Seq primer: AGCGGATAACAAATTTTCACACAGGA  
High quality sequence stop: 563.

FEATURES	SOURCE
1. <b>Feature 1</b>	Source 1
2. <b>Feature 2</b>	Source 2
3. <b>Feature 3</b>	Source 3
4. <b>Feature 4</b>	Source 4
5. <b>Feature 5</b>	Source 5
6. <b>Feature 6</b>	Source 6
7. <b>Feature 7</b>	Source 7
8. <b>Feature 8</b>	Source 8
9. <b>Feature 9</b>	Source 9
10. <b>Feature 10</b>	Source 10
11. <b>Feature 11</b>	Source 11
12. <b>Feature 12</b>	Source 12
13. <b>Feature 13</b>	Source 13
14. <b>Feature 14</b>	Source 14
15. <b>Feature 15</b>	Source 15
16. <b>Feature 16</b>	Source 16
17. <b>Feature 17</b>	Source 17
18. <b>Feature 18</b>	Source 18
19. <b>Feature 19</b>	Source 19
20. <b>Feature 20</b>	Source 20
21. <b>Feature 21</b>	Source 21
22. <b>Feature 22</b>	Source 22
23. <b>Feature 23</b>	Source 23
24. <b>Feature 24</b>	Source 24
25. <b>Feature 25</b>	Source 25
26. <b>Feature 26</b>	Source 26
27. <b>Feature 27</b>	Source 27
28. <b>Feature 28</b>	Source 28
29. <b>Feature 29</b>	Source 29
30. <b>Feature 30</b>	Source 30
31. <b>Feature 31</b>	Source 31
32. <b>Feature 32</b>	Source 32
33. <b>Feature 33</b>	Source 33
34. <b>Feature 34</b>	Source 34
35. <b>Feature 35</b>	Source 35
36. <b>Feature 36</b>	Source 36
37. <b>Feature 37</b>	Source 37
38. <b>Feature 38</b>	Source 38
39. <b>Feature 39</b>	Source 39
40. <b>Feature 40</b>	Source 40
41. <b>Feature 41</b>	Source 41
42. <b>Feature 42</b>	Source 42
43. <b>Feature 43</b>	Source 43
44. <b>Feature 44</b>	Source 44
45. <b>Feature 45</b>	Source 45
46. <b>Feature 46</b>	Source 46
47. <b>Feature 47</b>	Source 47
48. <b>Feature 48</b>	Source 48
49. <b>Feature 49</b>	Source 49
50. <b>Feature 50</b>	Source 50
51. <b>Feature 51</b>	Source 51
52. <b>Feature 52</b>	Source 52
53. <b>Feature 53</b>	Source 53
54. <b>Feature 54</b>	Source 54
55. <b>Feature 55</b>	Source 55
56. <b>Feature 56</b>	Source 56
57. <b>Feature 57</b>	Source 57
58. <b>Feature 58</b>	Source 58
59. <b>Feature 59</b>	Source 59
60. <b>Feature 60</b>	Source 60
61. <b>Feature 61</b>	Source 61
62. <b>Feature 62</b>	Source 62
63. <b>Feature 63</b>	Source 63
64. <b>Feature 64</b>	Source 64
65. <b>Feature 65</b>	Source 65
66. <b>Feature 66</b>	Source 66
67. <b>Feature 67</b>	Source 67
68. <b>Feature 68</b>	Source 68
69. <b>Feature 69</b>	Source 69
70. <b>Feature 70</b>	Source 70
71. <b>Feature 71</b>	Source 71
72. <b>Feature 72</b>	Source 72
73. <b>Feature 73</b>	Source 73
74. <b>Feature 74</b>	Source 74
75. <b>Feature 75</b>	Source 75
76. <b>Feature 76</b>	Source 76
77. <b>Feature 77</b>	Source 77
78. <b>Feature 78</b>	Source 78
79. <b>Feature 79</b>	Source 79
80. <b>Feature 80</b>	Source 80
81. <b>Feature 81</b>	Source 81
82. <b>Feature 82</b>	Source 82
83. <b>Feature 83</b>	Source 83
84. <b>Feature 84</b>	Source 84
85. <b>Feature 85</b>	Source 85
86. <b>Feature 86</b>	Source 86
87. <b>Feature 87</b>	Source 87
88. <b>Feature 88</b>	Source 88
89. <b>Feature 89</b>	Source 89
90. <b>Feature 90</b>	Source 90
91. <b>Feature 91</b>	Source 91
92. <b>Feature 92</b>	Source 92
93. <b>Feature 93</b>	Source 93
94. <b>Feature 94</b>	Source 94
95. <b>Feature 95</b>	Source 95
96. <b>Feature 96</b>	Source 96
97. <b>Feature 97</b>	Source 97
98. <b>Feature 98</b>	Source 98
99. <b>Feature 99</b>	Source 99
100. <b>Feature 100</b>	Source 100

```

1. .563
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="BP250023A10F2"
/clone_lib="Soares normalized bovine placenta"
/sex="female"
/lab_host="DH10B"
/note="Organ: placenta; Vector: pT73pac; Site 1: EcoRI;
Site 2: NotI; The cDNA library was constructed by the
Soares laboratory and it was constructed and normalized
as described by Bonaldo, M.F., Lennon, G. and Soares,
M.B. (1996), Genome Research 6(9): 791-806. "
110 a 177 c 102 t
BASE COUNT

```

BASE COUNT	110 a	177 c	174 g	102 t
ORIGIN				
Query Match	19.0%; Score 271.4; DB 12; Length 553;			
Best Local Similarity	96.0%; Pred. No. 2.4e-43;			

QY	1125	GCTGGGCTGGAGACACTGGATTATTCGGCCCGCTGGAGTACGAGGCTATCACTGGGAGGG	1184
Db	563	GCTGGGCTGGAGACACTGGATTATTCGGCCCGCTGGAGTACGAGGCTACCACTGCGAGGG	504
QY	1185	TGTATGCGACTTCCCGCTGGCGTCGCACTCGAGCGCCACCAACACGCGCATCATCCAGAC	1244
Db	503	CGTGTGCGACTTCCCGCTACGCTCGCACTGGAGCCCAACCAACGCGCATCATCCAGAC	444
QY	1245	GCTGATGAATCCATGAGACCCCGGCTCCACCCCGCCAGACTGCTGCTGGCGCCACCAAAATT	1304
Db	443	GCTGATGAATCCATGAGACCCCGGCTCCACCCCGCCAGCTGCTGCGTGGCGCCACCAAAATT	384
QY	1305	GACTCCCATCAGCAATTCATACATCGACGCGGGCAATAATGTGTCTCAACAAGCAGTAGCA	1364
Db	383	GACTCCCATCAGCATCTTGATCATCGACGCGGGCAATAATGTGTCTACAGCAGTAGCA	324
QY	1365	GGACATGGTGGTAGTCGTGCGGCTCGAGGTACGGTGCCTTTCCGCCGCCCTTGGGCC	1424
Db	323	GGAGATGGTGGTAGTCGTGCGGCTCGAGGTACGGGCGCC-TTCCCTGCCGCCCTCGGGCC	265
QY	1425	G 1425	
		↑	
Db	264	G 264	

RESULT 5	BQ169372	661 bp	linear	EST 01-MAY-2002
LOCUS	BQ169372	STR00320	segmentation stage cDNA library	Danio rerio cDNA clone
DEFINITION		CB278.5'	similar to DYNAMO PROTEIN, mRNA sequence.	
ACCESSION	BQ169372			
VERSION	BQ169372.1	GI:20376800		
KEYWORDS	EST.			
SOURCE				
ORGANISM				
				zebrafish.
				Danio rerio

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
AUTHORS	1 (bases 1 to 661) Loppin,B., Pflumio,S., Steffan,T., Meyer,V., Furthauer,M., Thisse,C. and Thisse,B.
TITLE	Expression of the zebrafish genome during embryogenesis (2002)
JOURNAL	Unpublished (2002)
COMMENT	Other_ESTs: STR00321 Contact: Thisse B Institut de Genetique et de Biologie Moleculaire et Cellulaire CNRS, INSERM, ULP 1, rue Laurent Fries, BP163, CU de Strasbourg, 67404 Illkirch Cedex , France Tel: 33 3 88 65 33 60 Fax: 33 3 88 65 32 01

Email: [thisse@igbmc.u-strasbg.fr](mailto:thisse@igbmc.u-strasbg.fr)  
EST from a cDNA of a gene whose expression is spatially restricted during embryogenesis. We have established its expression pattern during embryonic development by whole mount in situ hybridization on zebrafish embryos from the gastrula stage to 2 days of development. The corresponding data are available on the zebrafish community database at <http://zfinfo.org/> cDNA library preparation: B. Riddleman. DNA Sequencing by: IGBMC sequencing facility. Clone distribution: zebrafish international resource center at the University of Oregon (Institute of Neuroscience, 1254 University of Oregon, Eugene, OR 97403-1254)  
Seq primer: T3 ATTTACCCCTCACTCAATGAAGGGA.

FEATURES  
source

```

source
I. .661
location/Vaduz11r1e1s
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="CB278"
/clone_lib="segmentation stage cDNA library"
/dev_stage="segmentation stage embryos"
/note=Vector: lambda zap; Site_1: EcoRI; Site_2: XhoI;
Oligo dt cDNA library constructed from RNA pooled from
segmentation stage zebrafish embryos"
201 a 169 c 172 g 119 t
BASE COUNT
```

BASE COUNT	201 a	189 c	172 g	119 t
ORIGIN				
Query Match		17.8%	Score 253.6;	DB 14; Length 661;

QY	1005	GCTGCCCTCGCCCGCGCGCGCGCGCGCGCTTCGCCAGTGCCCATGCAACGC	1064
Db	198	GCTCAGCTTTAAAGCGCGGCGACAGAGAACTGCGCTTAATTAATCGGCACGGAAAAG	257
QY	1065	GCACGGCAAGAACTCCAGGCTACGCTGCAGCAAGAACCCCTGCACGTGAACTTCAAGGA	1124
Db	258	GCATGGCAAAAGTCCAAATTCGAGATCAGCAAAAGAGCTCTGCACGTCAACTTTCAAGA	317
QY	1125	GCTGGGCTGGAGCAGCTGGATTATTCGGCGCCCTCGAGTACGAGCGCTATCACTCGGAGGG	1184
Db	318	ACTTGGATGGAGCAGCTGGGATCATCGTCCCTCGATTACGAACCCATCACTCGGAGGG	377
QY	1185	TGTAATGCGACTTCOCGCTCGCGCTGCACCTCGAGCCCAACCAACCGCATCACTCCAGAC	1244
Db	378	CGTGTGCGACTTCOCGTTGAGTTCGACCTTAGAGCCGACCAACCAACCGCATCACTCCAGAC	437
QY	1245	GCTGATGAATCCTCATGAGACCCCGGCTCCACCCCGCCAGCTGCTGGTGGCCCAACCAAT	1304
Db	438	GCTCATGAATCCCATGAGCCCAACACGACATCCACCGAGCTGTTGCGTCCCAACCAACT	497
QY	1305	GACTTCCCATCAGCAATTCATACATCGACGGGGCAATAATGTGGTCTCTACAGCAGTAGCA	1364
Db	498	CAGCCCCATCAGTATATCTGATAGACTCTGGGAACACAGCTCGTGTACAAACAGTAGCA	557
QY	1365	GGACATGGTGGTGGAGTCGTGCGGCTCGAGGTAGCG	1400
Db	558	GGACATGGTGGTGAACACAGTGTGGCTCGAGTAGCG	593



TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

COMMENT Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msoares@blue.weeg.uiowa.edu  
CDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com) This clone is also available through the I.M.A.G.E. Consortium at LLNL (info@image.llnl.gov). IMAGE ID= 1768988  
Seq primer: M13 Forward.

# FEATURES

source  
1. .343  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-C3-sq-b-01-0-UI"  
/clone\_lib="UI-R-C3"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; The UI-R-C3 library is a subtracted library of a series, ultimately derived from a mixture of individually tagged normalized libraries from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, muscle, and 8, 12 and 18-day embryos, after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: UI-R-C3, UI-R-C2p, UI-R-C1, UI-R-C0, UI-R-A1, UI-R-E1. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within themixture. The subtracted library (UI-R-C3) was constructed as follows: PCR amplified cDNA inserts from UI-R-C2p clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-C2p library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C3 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)"  
Research 6:791-806, 1996"

BASE COUNT 63 a 99 c 108 g 72 t 1 others

Query Match 17.6%; Score 250.2; DB 12; Length 343;  
Best Local Similarity 83.9%; Pred. No. 3.3e-39;  
Matches 282; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 439 GATCTCTCGCACACTCTCTCCGGAGACAGAGTATTGTTGATGTGCTCATGCTCA 498  
|||||  
DB 8 GATCTCTCGCACACTCTCTCCGGAGACAGAGTATTGTTGATGTGCTCATGCTCA 67  
|||||

QY 499 GACAAAGAGAGTGTTGGCGCGGAGCTCGGCTCTTTTCGCCAGCGCCCTCAGGCGCC 558  
|||||  
DB 68 GACAAAGAGAGTGTTGGCGCGGAGCTCGGCTCTTTTCGCCAGCGCCCTCAGGCGCC 127  
|||||

QY 559 TGGGGGCCACACGGCGGCTCCAGTGCAGCTCTTCCTTCCTTCCTTCCTTCCTTCCT 618  
|||||

DB 128 TGGGGGCCACAAACCCGACGCTGGATTGTCAGCTCTTCCTTCCTTCCTTCCTTCCT 187  
|||||

QY 619 CTGGAGCGGGGACCTTGGACCGCGAGGGCGCGCGCGCTGGGAAGTCTTCGAC 678  
|||||

DB 188 CTGGACTCTAGGACCTTGGATCTCAGGGGCCAACAGAGCTGCTGGGAAGTCTTCNAC 247  
|||||

QY 679 GTGTGGCAGGCGCTCGCCACCACCGCTTGAAGCAGCTGTCTTGGAGCTCGGGCGGCA 738  
|||||

DB 248 GTGTGGCAGGCTCTCGCCCTCAGCCGTGAAGCAGCTGTCTTGGAGCTCGGGCAGTC 307  
|||||

QY 739 TGGGCGCAGCTGACGCCGGGGGAGCGCGAGCGCGC 774  
|||||

DB 308 TGGGGTGAGTGGACGCCAGGGATTACAGAGCGCGC 343  
|||||

RESULT 7  
BQ604531  
LOCUS BQ604531  
DEFINITION MI-P-CP1-nzc-d-04-0-UI.s1 MI-P-CP1 Sus scrofa cDNA clone  
ACCESSION BQ604531  
VERSION BQ604531.1 GI:21551257  
KEYWORDS EST.  
SOURCE Sus scrofa  
ORGANISM pig.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
REFERENCE 1 (bases 1 to 798)  
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
COMMENT Contact: Tuggle CK  
Molecular Genetics Laboratory, Department of Animal Science  
Iowa State University  
201 Kildee Hall, Ames, IA 50011-3150, USA  
Tel: 5152944252  
Fax: 5152942401  
Email: ktuggle@iastate.edu  
Tissue Procurement: Dr. Chris Tuggle, Iowa State University  
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).  
The following repetitive elements were found in this cDNA  
sequence: 1-86, >GC-rich#Low\_complexity (matched complement)  
Seq primer: M13 FORWARD  
POLYA=No.

# FEATURES

source  
1. .798  
/organism="Sus scrofa"  
/strain="crossbred"  
/db\_xref="taxon:9823"  
/clone="MI-P-CP1-nzc-d-04-0-UI"  
/clone\_lib="MI-P-CP1"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: EcoRI; The MI-P-CP1 library is normalized library derived from the MI-P-CP0 library, ultimately derived from uterus tissue. For a detailed description of the library from which this clone was derived, please visit our web site at http://pigest.genome.iastate.edu/. The procedure used to create this library has been previously described (Bonaldo , Lennon and Soares, Genome Research 6: 791-806, 1996)  
TAG\_SEQ=None found"  
BASE COUNT 169 a 227 c 265 g 137 t  
ORIGIN

Query Match 16.7%; Score 237.6; DB 14; Length 798;  
Best Local Similarity 74.3%; Pred. No. 1.1e-36;  
Matches 300; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 996 CAGGCCTTGGCTCGCCCGCGCGCGCGCGCGCGCTTCGCCAGTCCGCCA 1055  
|||||

DB 8 CAGGCGGTGG 67  
|||||







```

Db 1 CGCAAGAACTGTTCACCTGAGATGATGACAGCTGGGTCTGACAGAGCTGCCGGA--- 57
QY 946 GGCAGCGGGCGGCGAGGGGTCTGGCCGCGCGCGTTCGGGCGCCCGCGATGCCAGGCGCTTGG 1005
Db 58 -----GCCGAGGGGTTCATGGCCAGCGCGGTTCGGGCGCCCGACGCGCGGTCTTGG 108
QY 1006 CTGCGCTCGCCGCGCGCGCGGGCGGCGAGCGCCCTTCGCGAGTCGCCATGTCGAAGCGG 1065
Db 109 CTGCGCTCGCCGCGCGCGGGCGGCGAGCGCCCTTCGCGAGTCGCCATGTCGAAGCGG 168
QY 1066 CAGCGCAAGAAGTCCAGGCTAGCTGCAGCAAGAAGCCCTGCACGTGAACCTCAAGGAG 1125
Db 169 CAGCGCAAGAAGTCCAGGCTAGCTGCAGCAAGAAGCCCTGCACGTGAACCTCAAGGAG 228
QY 1126 CTGGGCTGGGACGACTGGATTATCGCGCCCGCTGGAGTAGAGGCGCTTACACTCGAGGGT 1185
Db 229 TTAGGCTGGGAGATTGGATTATCGCGCCCGCTAGAGTAGAGGCGCTTACACTCGAGGGC 288
QY 1186 G 1186
Db 289 G 289

RESULT 13
BQ563481/c
LOCUS
DEFINITION BQ563481 479 bp mRNA linear EST 19-JUN-2002
clone gi03h02.y1 Mouse Organ of Corti cDNA pBluescript Mus musculus cDNA
ACCESSION BQ563481
VERSION BQ563481.1 GI:21466450
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 479)
EST analysis of gene expression in the mouse Organ of Corti at the
onset of hearing
Unpublished (2002)
Contact: Kachar, B.
Structural Cell Biology
National Institute of Deafness and other Communication Disorders
50/4249 South Drive, NIH, Bethesda, MD 20892-8027, USA
Tel: 301-402-1599
Fax: 301-402-1765
Email: kacharbenidc@nih.gov
Plate: 03 row: h column: 02
Seq primer: M13RPI reverse primer (ABI).
Location/Qualifiers
1..479
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone_lib="gi03h02"
/sex="male and female"
/dev_stages="Post natal day 5 to 13"
/note="Organ: Organ of Corti; Vector: pBluescript; The
organ of Corti (OC) was fine dissected from a total of 386
OC as follows: 102 samples from post-natal (P) day 5; 72
from P6; 60 from P7; 46 from P8; 18 from P9; 20 from P10;
14 from P12 and 24 from P13. After killing animals by
cervical dislocation followed by decapitation, the bulla
was removed and opened in Leibowitz medium. The bony
capsule of the cochlea was chipped away, stria vascularis
and spiral ligament were removed and the sensory
epithelium was carefully dissected out of the modiolus.
Total RNA was extracted using the micro Fasttrack kit
(catalog # K1593-02; Invitrogen, Carlsbad, CA), according
to manufacturer's instructions. Reverse transcription and
library construction were carried out with the Uni-Zap XR
vector kit (catalog # 237211, Stratagene) and Uni-Zap XR

```

```

Gigapack III Gold Cloning kit (catalog # 237612), both
from Stratagene (La Jolla, CA, USA), according to
manufacturer's instructions. Briefly: 1.5 ug mRNA was
reverse transcribed using a hybrid oligo(dT) linker-primer
that contains an Xho I site. First strand synthesis was
primed with the linker-primer and transcribed using
Moloney murine leukemia virus reverse transcriptase
(MMLV-RT) and 5-methyl dCTP. The second strand was
synthesized with DNA polymerase and RNase H. Complementary
DNA was blunt ended with Pfu DNA polymerase, ligated with
EcoR I adapters in the presence of ligase and digested
with Xho I. The cDNA was sequentially size fractionated
over Pharmacia Size Sep400 (Pharmacia, Uppsala, Sweden)
and Clontech Chroma Spin-1000 (Clontech, Palo Alto, CA)
columns to enrich for cDNAs greater than 400bp and 1000 bp
, respectively. The cDNA was then directionally ligated to
the Uni-Zap XR vector, which had been predigested with
EcoR I and Xho I. The phagemid was packaged with Gigapak
III Gold and, upon titration on XL1 Blue MRF' cells, the
yield of the phage library was estimated to be 11,100,000
recombinants. Stratagene's EXAssist Interference
resistance helper phage (catalogue # 211203) was adopted
to rescue plasmid DNA from the phages. Upon plating of the
rescued library, individual cDNA clones were selected and
grown in 96-well, 2 ml growth plate. Plasmid DNA was
purified from 200 ul of saturated culture with the
Concert96(TM) plasmid purification kit (Invitrogen,
Carlsbad, CA) as instructed by the manufacturer. ESTs from
the 5' end of the cDNA clones were generated with the
universal M13 reverse primer (CAGGAACAGCTATGACC) and 25%
strength BigDye terminator sequencing chemistry (Applied
Biosystems, Foster City, CA). Sequencing reactions were
performed on MJ Tetrad thermal cyclers (MJ Research,
Waltham, MA), and analyzed on 3700 automated capillary
sequencers using POP5 polymer (Applied Biosystems, Foster
City, CA). The frequency distribution of the library is
as follows: 72% of genes have 1 copy; 14.3% 2; 12% 3-10;
1.4% 11-50 and 0.1% 51-150. As to gene function, 45% of
genes are present in GenBank and have known function; 23%
have hits in GenBank, but do not have assigned function;
12% are uncharacterized ESTs and 20% are unidentified."

BASE COUNT 105 a 119 c 144 g 111 t
ORIGIN
Query Match 13.5%; Score 191.8; DB 14; Length 479;
Best Local Similarity 74.6%; Pred. No. 9.3e-28;
Matches 241; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 1021 CGCCGCGCGCGCACGCGCCTTCGCCAGTGCAGCGGCGGCAAGAGATCC 1080
Db 324 CGCGCGAAACGCGCGGCCCATTTGGCCAAATCGCGAGGCAAGCCACGACGACCTC 265
QY 1081 AGGCTACGCTGCAGCAAGAGCCCTTCGAGTGAACCTTCAGGAGCTGGCTGGGACGAC 1140
Db 264 AGGCTCGCTGCAGTGCAGGCGCTTGCATGCAACTTCAAGGACATGGGCTGGGACGAC 205
QY 1141 TGGATTATCGCGCCCTCGGATGACGAGGCTATCACTGGGAGGCTGTATCGGACTTCCCG 1200
Db 204 TGGATCATCGACCTCTTGAGTATGAGGCTTCCACTGCGAAGAGACTGTGTGATTCCCC 145
QY 1201 CTGCGCTCGCACCTTGAGGCGCCCAACACCGGCGCATCATCGACGCTGATGAACCTCATG 1260
Db 144 TTGCGCTCCACCTTGAGGCGCCCAACACCGGCGCATCATCGACGCTGATGAACCTCATG 85
QY 1261 GACCCGCGCTCCACCCCGCGGCGCTGCTGGTGGCCCAACAAATGACTCCCATCAGCAT 1320
Db 84 GACCCCTGAATCCACACCGCCACTTGTGTGTGCTTACACGGCTGAGTCTCTATTAGCATC 25
QY 1321 CTATACATCGACGCGGCGCAATA 1343
Db 24 CTCTTCATCGACTCTGCCAACAA 2

```



Search completed: November 25, 2002, 02:50:44  
Job time : 2445 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 23:26:25 ; Search time 310 Seconds  
(without alignments)  
10351.928 Million cell updates/sec

Title: US-09-825-751A-19

Perfect score: 1425

Sequence: 1 ctctcgggagcagcgacca.....tttccgcgcgttgccgcg 1425

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002:\*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1425	100.0	1425	22	RAI70203 Human growth/diffe
2	1425	100.0	1425	24	ABK43993 DNA encoding human
3	1421.8	99.8	2791	24	AA28496 Human extracellular
4	987	69.3	1171	18	AAT90386 Human bone morphog
5	987	69.3	1171	22	AA28496 Human full length
6	976	68.5	1171	16	AAQ96208 Human mature VL-1
7	928.2	65.1	1308	17	AAT31602 Cartilage-derived
8	581.8	40.8	1046	18	AAT90397 Murine BMP-13 homo
9	581.8	40.8	1046	22	AA28496 Murine partial mv2

10	581.2	40.8	1046	16	AAQ96224 Murine mv2 protein
11	412.2	28.9	530	16	AAQ84280 GDF-6 gene. Mus s
12	412.2	28.9	530	21	AAA72621 Murine growth diff
13	259.2	18.2	1345	16	AAQ96215 Human bone morphog
14	259.2	18.2	1345	18	AAQ90390 Human bone morphog
15	259.2	18.2	1345	22	AA28496 Human bone morphog
16	253	17.8	2703	24	ABQ8241 Human osteoblast d
17	253	17.8	2703	16	AAQ83695 New TGF-beta famil
18	253	17.8	2703	18	AAT98191 cDNA for human MP5
19	253	17.8	2703	18	AAQ9695 cDNA encoding huma
20	253	17.8	2703	18	AAQ6412 cDNA encoding huma
21	253	17.8	2703	18	AAT59405 DNA encoding human
22	253	17.8	2703	18	AAT59729 Human bone morphog
23	253	17.8	2703	19	AAQ8340 Human MP52 cDNA.
24	253	17.8	2703	24	ABQ8242 Human osteoblast d
25	251.6	17.7	1207	14	AAQ47709 TGF-beta-like clon
26	251.6	17.7	1207	16	AAQ96209 Murine protein MP5
27	251.6	17.7	1207	22	AA28496 Human MP-52 DNA.
28	251.4	17.6	2341	17	AAQ18317 Cartilage-derived
29	250	17.5	2703	22	AAQ74420 Human TGF-beta MP5
30	249.2	17.5	519	16	AAQ82824 GDF-7 C-terminal r
31	245.8	17.2	357	17	AAQ46150 cDNA encoding huma
32	245.8	17.2	357	18	AAQ70296 Human bone inducin
33	245	17.2	1203	16	AAQ96223 Murine mv1 protein
34	245	17.2	1203	18	AAT90396 Murine BMP-12 homo
35	245	17.2	1203	22	AAQ18335 Murine partial mv1
36	244.8	17.2	926	16	AAQ96207 Human bone morphog
37	244.8	17.2	926	18	AAT90385 Human bone morphog
38	244.8	17.2	926	22	AA28496 Human bone morphog
39	239	16.8	2328	15	AAQ70010 Growth differentia
40	239	16.8	2329	22	AAH28140 Nucleotide sequenc
41	237.8	16.7	357	21	AAZ29328 Mutant human MP52
42	234.6	16.5	3598	23	AAQ93562 DNA encoding novel
43	232	16.3	272	16	AAQ96221 Murine bone morpho
44	232	16.3	272	22	AA28496 Murine clone, mv2
45	230	16.1	1233	16	AAQ96226 Fusion of BMP-2 pr

ALIGNMENTS

RESULT 1  
AAI70203  
ID AAI70203 standard; DNA; 1425 BP.

AC AAI70203;

DT 07-JAN-2002 (first entry)

Human growth/differentiation factor-6-like protein AMF10 DNA.

AMF10; human; growth/differentiation factor-6; cancer;  
cell proliferation; astrocytoma; glioma; gene therapy; diagnosis;  
ds.

OS Homo sapiens.

Key Location/Qualifiers

FT CDS 31..1398

FT sig\_peptide /\*tag= a

FT mat\_peptide /\*tag= b

FT 97..1395

FT /\*tag= c

PN WO200174897-A2.

XX 11-OCT-2001.

XX 03-APR-2001; 2001WO-US10892.

XX 03-APR-2000; 2000US-194314P.

PR 16-AUG-2000; 2000US-225693P.

XX (CURA-) CURAGEN CORP.  
XX Vernet CAM, Burgess CE, Fernandes E, Taupier RJ, Quinn KE;  
PI Spytek KA, Rastelli L, Herrmann JL;  
XX WPI; 2001-626395/72.  
DR P-PSDB; AAM50216.  
XX New AMF1-10 polypeptides and encoding polynucleotides, useful for  
PT treating or preventing disorders related to modulation of cell  
PT movement, cell signal processing, cell adhesion or migration pathways  
PT e.g., cancer  
XX Claim 3; Page 44-45; 134pp; English.  
XX The present sequence is that of DNA clone AC01136.A. This  
CC includes an open reading frame that encodes the novel, secreted  
CC human growth/differentiation factor-6 (GDF6)-like protein, AMF10  
CC (see AAM50216). AMF10 is expressed in astrocytoma and glioma-derived  
CC tissue. DNA encoding the AMF8 may be useful in gene therapy, and the  
CC protein may also be used as a therapeutic, especially in treatment of  
CC cancer and other cell proliferative disorders. Generally, the  
CC AMF1-10 (AMFX) nucleic acids and proteins of the invention are useful  
CC for treating or preventing AMFX-associated disorders, e.g. a disorder  
CC related to cell signal processing and metabolic pathway modulation,  
CC cell adhesion or migration pathway modulation, chemoresistance,  
CC radiotherapy resistance, survival in trophic factor limited  
CC secondary tissue site microenvironments, connective tissue  
CC disorders, tissue remodeling, oncogenesis, cancer of the breast,  
CC ovary, cervix, prostate, endometrium, stomach, colon, lung,  
CC bladder, kidney, brain, and soft-tissue, cellular transformation,  
CC developmental tissue remodeling, inflammation, blood clot  
CC formation and resorption, haematopoiesis, angiogenesis, multirug  
CC resistance related to organic anion transporters, malignant disease  
CC progression, autocrine and paracrine regulation of cell growth, and  
CC cellular responses to external stimuli, and other diseases,  
CC disorders, etc. (all claimed). Nucleic acid fragments are also  
CC useful as probes or PCR primers, for modulating AMFX expression,  
CC in chromosome mapping, tissue typing, diagnostics, forensics and  
CC pharmacogenomics.  
XX  
SQ Sequence 1425 BP; 234 A; 488 C; 469 G; 234 T; 0 other;  
Query Match 100.0%; Score 1425; DB 22; Length 1425;  
Best Local Similarity 100.0%; Pred. No. 6.7e-246;  
Matches 1425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTCCTGGGAGACGACGACCTTGGCCGCGCATGGATACCTCCAGGGTCTGCTCTCGGCC 60  
Db 1 CTCCTGGGAGACGACGACCTTGGCCGCGCATGGATACCTCCAGGGTCTGCTCTCGGCC 60  
QY 61 GTCCTTCCTCATCAGTTTCTGTGGGATTTGCCGGTTTCCAGCAGGCTTCCATCTCATCC 120  
Db 61 GTCCTTCCTCATCAGTTTCTGTGGGATTTGCCGGTTTCCAGCAGGCTTCCATCTCATCC 120  
QY 121 TCCTGTCTGTCGCCGAGCTGGGTTCCACCAAGGGCATCGGAAGCGCGAAGGAAGCAAG 180  
Db 121 TCCTGTCTGTCGCCGAGCTGGGTTCCACCAAGGGCATCGGAAGCGCGAAGGAAGCAAG 180  
QY 181 ATGACAGCGGGCGCGCGGACAGTACGACGGCGGGGAGGGCCAGGAACACAGCCGCGGG 240  
Db 181 ATGACAGCGGGCGCGCGGACAGTACGACGGCGGGGAGGGCCAGGAACACAGCCGCGGG 240  
QY 241 CCTCAGGAGAACCCCGGGCTCAGCAGCCCGGGGCGCAGAGCCGCGAGGGGTCCG 300  
Db 241 CCTCAGGAGAACCCCGGGCTCAGCAGCCCGGGGCGCAGAGCCGCGAGGGGTCCG 300  
QY 301 CGCGTGGTCCCGACAGTACATGCTGTCAATCTACAGGACTTACTTCCATCTGCTGAGAG 360  
Db 301 CGCGTGGTCCCGACAGTACATGCTGTCAATCTACAGGACTTACTTCCATCTGCTGAGAG 360  
QY 361 CTGGGCATCAATGCCAGCTTTTTCAGTCTTCCAAAGTCGGCTTAATACGATCACCAGCTTT 420

Db 361 CTGGGCATCAATGCCAGCTTTTTCAGTCTTCCAAAGTCGGCTTAATACGATCACCAGCTTT 420  
QY 421 GTAGACAGGGGACTAGACGATCTCGCACACTCTCTCCGGAGACAGAAATTTGTTT 480  
Db 421 GTAGACAGGGGACTAGACGATCTCTCGCACACTCTCTCCGGAGACAGAAATTTGTTT 480  
QY 481 GATGTGTCCATCTCTCAGACAAAGAAGAGCTGTGGGCGGAGCTGGGCTCTTTCCG 540  
Db 481 GATGTGTCCATCTCTCAGACAAAGAAGAGCTGTGTGGGCGGAGCTGGGCTCTTTCCG 540  
QY 541 CAGGCGCCCTCAGCGCCCTGGGGGCCACAGCCGGGCGGCTCCACGTGAGCTCTTCCCT 600  
Db 541 CAGGCGCCCTCAGCGCCCTGGGGGCCACAGCCGGGCGGCTCCACGTGAGCTCTTCCCT 600  
QY 601 TGCCTTTGCCCTTACTGCTGACGCGGGAGCCCTTGGACCCGAGGGGCGCGCGGCC 560  
Db 601 TGCCTTTGCCCTTACTGCTGACGCGGGAGCCCTTGGACCCGAGGGGCGCGCGGCC 560  
QY 661 GGCTGGGAAGTCTTCGACGTGTGGAGGCGCTGCGCCACACAGCCCTTGAAGCAGCTGTGC 720  
Db 661 GGCTGGGAAGTCTTCGACGTGTGGAGGCGCTGCGCCACACAGCCCTTGAAGCAGCTGTGC 720  
QY 721 TTGGAGCTGCGGGCGCATTTGGGCGAGCTTGGACGCCGGGAGGCGGAGCGCGCGGG 780  
Db 721 TTGGAGCTGCGGGCGCATTTGGGCGAGCTTGGACGCCGGGAGGCGGAGCGCGCGGG 780  
QY 781 GSACCCCAACACCGCCCGCCCGGACCTTGGCGAGTCTGGGCTTGGCGGAGGCTGCGG 840  
Db 781 GSACCCCAACACCGCCCGCCCGGACCTTGGCGAGTCTGGGCTTGGCGGAGGCTGCGG 840  
QY 841 CTCTCCAGAGCGGGCCCTGCTGTGTGTTATTCACAGATCCACGCGCAAGAACCTGTTC 900  
Db 841 CTCTCCAGAGCGGGCCCTGCTGTGTGTTATTCACAGATCCACGCGCAAGAACCTGTTC 900  
QY 901 GCAGAGATGCGGAGCAGCTGGGCTGCGGCCAGGCTGCGGGCCCGGGCGCGCGCGGAG 960  
Db 901 GCAGAGATGCGGAGCAGCTGGGCTGCGGCCAGGCTGCGGGCCCGGGCGCGCGCGGAG 960  
QY 961 GGGTGTGGCCCGCCCGCTGCGGGCCCGCGGATGCCAGGCTTGGCTGCCCTCGCCCGGC 1020  
Db 961 GGGTGTGGCCCGCCCGCTGCGGGCCCGCGGATGCCAGGCTTGGCTGCCCTCGCCCGGC 1020  
QY 1021 CGCCGGCGCGCGCACGGCCCTTCGCCAGTCCGCAATGCAAGCGGACGCGCAAGATCC 1080  
Db 1021 CGCCGGCGCGCGCACGGCCCTTCGCCAGTCCGCAATGCAAGCGGACGCGCAAGATCC 1080  
QY 1081 AGGCTACGCTGCAGCAAGAAGCCCTTGACGCTGAACCTTCAAGGAGCTGGGCTGGACGAC 1140  
Db 1081 AGGCTACGCTGCAGCAAGAAGCCCTTGACGCTGAACCTTCAAGGAGCTGGGCTGGACGAC 1140  
QY 1141 TGGATTATCGCCCGCTGGAGTACGAGGCTATCACTGCGAGGGTGTATGCGACTTCCG 1200  
Db 1141 TGGATTATCGCCCGCTGGAGTACGAGGCTATCACTGCGAGGGTGTATGCGACTTCCG 1200  
QY 1201 CTGCGCTGCGACCTGGAGCCCAACCAACGCGCATCATTCAGACGCTGATGAATCCATG 1260  
Db 1201 CTGCGCTGCGACCTGGAGCCCAACCAACGCGCATCATTCAGACGCTGATGAATCCATG 1260  
QY 1261 GACCCCGGCTCCACCCCGCCAGCTGTGCTGCGCCACCAAAATGACTCCCATACGACTT 1320  
Db 1261 GACCCCGGCTCCACCCCGCCAGCTGTGCTGCGCCACCAAAATGACTCCCATACGACTT 1320  
QY 1321 CTATACATCGAGCGGCAATATGCTGTCTACAGCAGTACGAGGACATGCTGTGGAG 1380  
Db 1321 CTATACATCGAGCGGCAATATGCTGTCTACAGCAGTACGAGGACATGCTGTGGAG 1380  
QY 1381 TCGTCCGCTGCGAGTACGCTGCTTTCCCGCGCCCTTGGCCG 1425  
Db 1381 TCGTCCGCTGCGAGTACGCTGCTTTCCCGCGCCCTTGGCCG 1425

RESULT 2

ABK48393  
ID ABK48393 standard; DNA; 1425 BP.  
AC  
XX ABK48393;  
XX  
XX 02-JUL-2002 (first entry)  
XX  
DE DNA encoding human growth/differentiation factor 6-like protein NOV9.  
XX  
XX Human; growth/differentiation factor 6-like protein; NOV9; gene;  
KW vaccine; cancer; tumour; bone disorder; avascular necrosis; allergy;  
KW haematopoietic disorder; immune disorder; endometriosis; renal disease;  
KW infection; inflammatory disease; lung disease; scleroderma; ataxia;  
KW bowel disease; appendicitis; blood disorder; cardiovascular disorder;  
KW graft versus host disease; GVHD; lymphoedema; brain disorder;  
KW ocular disorder; hepatitis C virus infection; cardiac disorder;  
KW autosomal dominant deafness; DFNA-2; ds.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH 7..1254  
CDS /\*tag= a  
FT /product= "Human growth/differentiation factor 6-like  
FT protein NOV9"  
FT  
XX WO200214368-A2.  
XX  
XX 21-FEB-2002.  
XX  
XX 16-AUG-2001; 2001WO-US25624.  
XX  
XX 16-AUG-2000; 2000US-225692P.  
XX 16-AUG-2000; 2000US-225693P.  
XX 16-AUG-2000; 2000US-225837P.  
XX 18-AUG-2000; 2000US-226236P.  
XX 18-AUG-2000; 2000US-226353P.  
XX 22-AUG-2000; 2000US-227085P.  
XX 23-AUG-2000; 2000US-227395P.  
XX 24-AUG-2000; 2000US-227492P.  
XX 24-AUG-2000; 2000US-227600P.  
XX 14-MAR-2001; 2001US-275952P.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Zerhusen BD, Padigar M, Spytek KA, Spaderna SK, Gangolli EA;  
PI Rastelli L, Burgess CE, Majumder K, Shimkets R, Mishra V;  
PI Vernet CAM, Szekeres ES, Grosse WM, Alsobrook JP, Liu X;  
PI Gerlach VL, Ellerman K, Smithson G, Peyman J, Stone D;  
PI MacDougall J;  
XX  
XX WPI; 2002-329571/36.  
DR P-PSDB; AAU79173.  
DR  
XX  
XX Novel cytoplasmic, nuclear membrane bound and secreted NOVX  
PT polypeptides, useful for treating cancers and tumours, bone disorders,  
PT Paget's disease, haematopoietic disorders, spinal diseases and immune  
PT disorders -  
XX  
XX Claim 8; Page 86; 234pp; English.  
XX  
XX The present invention relates to new isolated NOVX polypeptides named  
CC NOV1-NOV9. The invention can be used for identifying an agent (a cellular  
CC receptor or downstream effector) that binds to the polypeptide. The  
CC molecules of the invention are useful for treating or preventing  
CC NOVX-associated disorders in humans. The antibody of the invention is  
CC useful for determining the presence or amount of NOVX in a sample, and  
CC for treating a pathological state in a mammal. The method of the  
CC invention is useful for determining the presence of an amount of NOVX in  
CC a sample which is used as a marker for cancerous cell or tissue type.  
CC The molecules of the invention are useful in the manufacture of a  
CC medicament for treating or preventing cancer, tumour, bone disorders,  
CC avascular necrosis, allergy, haematopoietic disorders, immune disorders,

CC endometriosis, renal diseases, infections, inflammatory diseases, lung  
CC diseases, scleroderma, ataxia, bowel diseases, appendicitis, blood  
CC disorders, cardiovascular disorders, graft versus host disease (GVHD),  
CC lymphoedema, brain disorders, ocular disorders, hepatitis C virus  
CC infection, cardiac disorders and autosomal dominant deafness (DFNA-2).  
CC The present nucleic acid sequence encodes the human  
CC growth/differentiation factor 6-like protein NOV9 that is one of the  
CC NOVX proteins described in the invention.  
XX  
XX Sequence 1425 BP; 234 A; 488 C; 469 G; 234 T; 0 other;  
Query Match 100.0%; Score 1425; DB 24; Length 1425;  
Best Local Similarity 100.0%; Pred. No. 6.7e-246;  
Matches 1425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTCCTGGGAGAGCGACGACCTTGCCTCCATGGATCTCCAGGCTCTGCTCTCGGCC 60  
DB 1 CTCCTGGGAGAGCGACGACCTTGCCTCCATGGATCTCCAGGCTCTGCTCTCGGCC 60  
QY 61 GTCCTCTCATCAGTTTCTGTGGGATTTGCCGGTTTCCAGCAGGCTTCCATCTCATCC 120  
DB 61 GTCCTCTCATCAGTTTCTGTGGGATTTGCCGGTTTCCAGCAGGCTTCCATCTCATCC 120  
QY 121 TCGTGTTCGTCGCGAGCTGGGTTCCACCAAGGGCATCGGAAGCGCAAGAGGCAAG 180  
DB 121 TCGTGTTCGTCGCGAGCTGGGTTCCACCAAGGGCATCGGAAGCGCAAGAGGCAAG 180  
QY 181 ATGCAGCGGGCGCGCGGACAGTACGCGGGCGGAGGCGCAGCAACAGCGCGG 240  
DB 181 ATGCAGCGGGCGCGCGGACAGTACGCGGGCGGAGGCGCAGCAACAGCGCGG 240  
QY 241 CCTCAGGAGCAACCCCGGGCTCAGCAGCGCCCGGCGCAGGCGCGGAGGGTCCG 300  
DB 241 CCTCAGGAGCAACCCCGGGCTCAGCAGCGCCCGGCGCAGGCGCGGAGGGTCCG 300  
QY 301 CGCGTGTGTCGCGCGAGTACATGTGTCAATCTACAGGACTTACTTCCATCGTGAAG 360  
DB 301 CGCGTGTGTCGCGCGAGTACATGTGTCAATCTACAGGACTTACTTCCATCGTGAAG 360  
QY 361 CTGGGCATCAATGCAGCTTTTTCAGTCTTCCAGTGGCTTAATACGATCACCAGCTTT 420  
DB 361 CTGGGCATCAATGCAGCTTTTTCAGTCTTCCAGTGGCTTAATACGATCACCAGCTTT 420  
QY 421 GTAGACGGGGACTAGACGATCTCTCGCACACTCTCTCCGGAGACAGATTTGTTT 480  
DB 421 GTAGACGGGGACTAGACGATCTCTCGCACACTCTCTCCGGAGACAGATTTGTTT 480  
QY 481 GATGTGTCATGCTCTCAGACAAAGAGCTGTGGGGCGGAGCTGCGGCTCTTTTCG 540  
DB 481 GATGTGTCATGCTCTCAGACAAAGAGCTGTGGGGCGGAGCTGCGGCTCTTTTCG 540  
QY 541 CAGGCGCCCTCAGGCGCCCTGGGGCCACAGCGGGCGGCTCCAGTGCAGCTCTTCCT 600  
DB 541 CAGGCGCCCTCAGGCGCCCTGGGGCCACAGCGGGCGGCTCCAGTGCAGCTCTTCCT 600  
QY 601 TGCCTTTCGCCCTACTGCTGGACGCGGAGCCCTGGACCCGCGGGGCGCGCGGCC 660  
DB 601 TGCCTTTCGCCCTACTGCTGGACGCGGAGCCCTGGACCCGCGGGGCGCGCGGCC 660  
QY 661 GSGTGGGAAGTCTTCGACGTGTGGCAGGGCCCTGCGCCACAGCCCTGGAAGCAGTGTGC 720  
DB 661 GSGTGGGAAGTCTTCGACGTGTGGCAGGGCCCTGCGCCACAGCCCTGGAAGCAGTGTGC 720  
QY 721 TTGGAGTCTCGGGCGCATGGGCGAGCTGGACGCGGGGAGGCGGAGCGCGCGCGG 780  
DB 721 TTGGAGTCTCGGGCGCATGGGCGAGCTGGACGCGGGGAGGCGGAGCGCGCGCGG 780  
QY 781 GGACCCCAACCGCGCCCGGACCTTCGCGAGTCTGGGCTTCGGCGGAGGTTGCGG 840  
DB 781 GGACCCCAACCGCGCCCGGACCTTCGCGAGTCTGGGCTTCGGCGGAGGTTGCGG 840  
QY 841 CCTCCCGCAGGCGGCGCTCTGTTGTTTACACAGATCCAGCGCAAGAACCTGTTTC 900  
DB 841 CCTCCCGCAGGCGGCGCTCTGTTGTTTACACAGATCCAGCGCAAGAACCTGTTTC



```
QY 481 GATGTGTCCATGCTCTCAGACAAAGAGAGCTGGTGGGCGGAGCTGGGCTCTTTTCGC 540
|
|
|
Db 1479 GATGTGTCCATGCTCTCAGACAAAGAGAGCTGGTGGGCGGAGCTGGGCTCTTTTCGC 1538
|
|
|
QY 541 CAGGCGCCCTACAGCGCCCTGGGGGCGACACAGCGGGCGCGTCCACGTGAGCTCTTCCCT 600
|
|
|
Db 1539 CAGGCGCCCTACAGCGCCCTGGGGGCGACACAGCGGGCGCGTCCACGTGAGCTCTTCCCT 1598
|
|
|
QY 601 TGGCTTTTGGCCCTACTGCTGGACGCGCGACCTCGACCCGAGGGGCGCGCGCGCG 660
|
|
|
Db 1599 TGGCTTTTGGCCCTACTGCTGGACGCGCGACCTCGACCCGAGGGGCGCGCGCGCG 1658
|
|
|
QY 661 GGTGGGAAGTCTTTCGACGCTGTGGCAGGGCGCTGCGCCACACAGCCCTGGAAGCAGCTGTGC 720
|
|
|
Db 1659 GGTGGGAAGTCTTTCGACGCTGTGGCAGGGCGCTGCGCCACACAGCCCTGGAAGCAGCTGTGC 1718
|
|
|
QY 721 TTGGAGCTGGGGCGCGATGGGGCGGAGCTGGAGCGCGGGGAGCGCGCGCGCG 780
|
|
|
Db 1719 TTGGAGCTGGGGCGCGATGGGGCGGAGCTGGAGCGCGGGGAGCGCGCGCGCG 1778
|
|
|
QY 781 GGACCCAGCAACCGCGCGCCCGGACCTGCGGAGTCTGGGCTTTCGCGGAGGGTGGCG 840
|
|
|
Db 1779 GGACCCAGCAACCGCGCGCCCGGACCTGCGGAGTCTGGGCTTTCGCGGAGGGTGGCG 1838
|
|
|
QY 841 CTTCCCGAGGAGCGGCGCTGCTGGTGGTATTCACAGATCCCGAGCGCAAGAACCTGTTC 900
|
|
|
Db 1839 CTTCCCGAGGAGCGGCGCTGCTGGTGGTATTCACAGATCCCGAGCGCAAGAACCTGTTC 1898
|
|
|
QY 901 GCAGATGCGCGAGCAGCTGGGCTCGCGGAGGCTGGGGCGCGGGCGGGCGCGGAG 960
|
|
|
Db 1899 GCAGATGCGCGAGCAGCTGGGCTCGCGGAGGCTGGGGCGCGGGCGGGCGCGGAG 1958
|
|
|
QY 961 GGTCTGTCGCGCGCGCTGCGGCGCCCGGATGCCAGGCTTGGCTTTCGCGCGCGCG 1020
|
|
|
Db 1959 GGTCTGTCGCGCGCGCTGCGGCGCCCGGATGCCAGGCTTGGCTTTCGCGCGCGCG 2018
|
|
|
QY 1021 CGCGGCGGCGCGCACGGGCTTCGCGAGTTCGCGATGGCAAGCGGCACGCGCAAGATCC 1080
|
|
|
Db 2019 CGCGGCGGCGCGCACGGGCTTCGCGAGTTCGCGATGGCAAGCGGCACGCGCAAGATCC 2078
|
|
|
QY 1081 AGCTACGCTGCAGCAAGAACGCCCTGCGAGTGAACTTCAAGAGCTGGGCTGGGACGAC 1140
|
|
|
Db 2079 AGCTACGCTGCAGCAAGAACGCCCTGCGAGTGAACTTCAAGAGCTGGGCTGGGACGAC 2138
|
|
|
QY 1141 TGGATTATCGCGCCCTGGAGTACGAGGCTATCACTGGGAGGCTGTATGCGACTTCCCG 1200
|
|
|
Db 2139 TGGATTATCGCGCCCTGGAGTACGAGGCTATCACTGGGAGGCTGTATGCGACTTCCCG 2198
|
|
|
QY 1201 CTGCGTGGACCTGGAGGCCCAACACCGGCATCATCCAGACGCTGATGAACCTCCATG 1260
|
|
|
Db 2199 CTGCGTGGACCTGGAGGCCCAACACCGGCATCATCCAGACGCTGATGAACCTCCATG 2258
|
|
|
QY 1261 GACCCGGCTCCACCGCGCCAGCTGCTGGTGGTCCCAACAAATTGACTCCCATCAGATT 1320
|
|
|
Db 2259 GACCCGGCTCCACCGCGCCAGCTGCTGGTGGTCCCAACAAATTGACTCCCATCAGATT 2318
|
|
|
QY 1321 CTATACATCAGCGCGGCAATAATGTGGTCTACAACAGTACGAGGACATGGTGGAG 1380
|
|
|
Db 2319 CTATACATCAGCGCGGCAATAATGTGGTCTACAACAGTACGAGGACATGGTGGAG 2378
|
|
|
QY 1381 TCGTGGGCTGCAGGTAGCGGTGCCCTTTCCCGCGCGCTTGGCCCG 1425
|
|
|
Db 2379 TCGTGGGCTGCAGGTAGCGGTGCCCTTTCCCGCGCGCTTGGCCCG 2423
|
|
|
```

RESULT 4  
AAT90386  
ID AAT90386 standard; DNA; 1171 BP.  
XX  
AC AAT90386;  
XX  
XX 21-JAN-1998 (first entry)

```
XX Human bone morphogenetic protein BMP-13 gene clone v1-1.
DE
XX
KW BMP-13; bone morphogenetic protein; human; tendon; ligament;
KW wound healing; tissue repair; tendonitis; carpal tunnel syndrome;
KW therapy; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 2..967
XX /tag= a
FT sig_peptide 2..604
FT /tag= b
FT mat_peptide 605..964
FT /tag= c
XX
XX US558882-A.
XX
XX 19-AUG-1997.
XX
XX 07-DEC-1993; 93US-0164103.
XX
XX 22-DEC-1994; 94US-0362670.
PR 07-DEC-1993; 93US-0164103.
PR 25-MAR-1994; 94US-0217780.
PR 02-NOV-1994; 94US-0333576.
XX
XX (GEMY ) GENETICS INST INC.
PA (HARD ) HARVARD COLLEGE.
XX
XX Celeste AJ, Melton DA, Rosen VA, Thomsen GH, Wolfman NM;
PI Wozney JM;
XX
XX WPI; 1997-424270/39.
DR P-PSDB; AAW26591.
XX
XX Inducing tendon and ligament formation using BMP-12, BMP-13 or MP-52
PT - useful for tissue healing and repair, treatment of tendonitis,
PT improving fixation of tendons to bone etc
XX
XX Example 1; Column 55-58; 43pp; English.
XX
XX This nucleotide sequence, designated v1-1, encodes a novel bone
XX morphogenetic protein (see AAW26591), designated BMP-13. It was
XX isolated from a human genomic library using a probe (see AAT90398)
XX based on a partial BMP-13 sequence. A claimed method for inducing
XX formation of tendon and/or ligament tissues involves administration
XX of a composition containing at least one protein selected from
XX BMP-13, MP52 (see AAW26590) and BMP-12 (see AAW26589). The method is
XX used for tissue (including skin) healing and repair. This is
XX useful for treating tendonitis, carpal tunnel syndrome and other
XX defects of traumatic or congenital origin, in cosmetic surgery and
XX to improve fixation of tendons or ligaments to bone. The specified
XX proteins can also be used to increase activity of other BMPs e.g.
XX BMP-2 (see AAW26597).
XX
XX Sequence 1171 BP; 195 A; 384 C; 413 G; 179 T; 0 other;
SQ
Query Match 69.3%; Score 987; DB 18; Length 1171;
Best Local Similarity 100.0%; Pred. NO. 1.2e-167;
Matches 987; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 439 GATCTCTCGCACACTCTCTCCGGAGACAGAGATTTCTTTGATGTGTCATGCTCTCA 498
|
|
|
Db 8 GATCTCTCGCACACTCTCTCCGGAGACAGAGATTTCTTTGATGTGTCATGCTCTCA 67
|
|
|
QY 499 GACAAAGAGAGCTGGTGGCGCGGAGCTGCGGCTCTTTTCGCGAGCGCGCTCAGGCGCC 558
|
|
|
Db 68 GACAAAGAGAGCTGGTGGCGCGGAGCTGCGGCTCTTTTCGCGAGCGCGCTCAGGCGCC 127
|
|
|
QY 559 TGGGGGCCACCGCGCGCGCTTCCACGCTTCCCTTGCCTTGCCTTGCCTTACTG 618
|
|
|
```

```
Db 128 TGGGGCCACACAGCGGGCGCTCCACGTGAGCTCTTCCCTTGCCCTTTCGCCCTACTG 187
QY 619 CTGGAGCGGGACCTTGACCGCAGGGGGCGCCGCCGGCTGGAAGTCTTCGAC 678
Db 188 CTGGAGCGGGACCTTGACCGCAGGGGGCGCCGCCGGCTGGAAGTCTTCGAC 247
QY 679 GTGTGGCAGGGCTGCGCCACACAGCCCTTGAAGCAGCTGTGCTTGAGCTGGGGCGCA 738
Db 248 GTGTGGCAGGGCTGCGCCACACAGCCCTTGAAGCAGCTGTGCTTGAGCTGGGGCGCA 307
QY 739 TGGGGCAGCTGGACCGCGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 798
Db 308 TGGGGCAGCTGGACCGCGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 367
QY 799 CCCCCGAGCTGGGAGTCTGGGCTTTCGGCCGAGGGTGGGGCTCCCGAGGCGGGCC 858
Db 368 CCCCCGAGCTGGGAGTCTGGGCTTTCGGCCGAGGGTGGGGCTCCCGAGGCGGGCC 427
QY 859 CTGCTGCTGTTATTCACAGATCCACAGCGCAAGAACTGTTCGACAGATGCGCAGCAG 918
Db 428 CTGCTGCTGTTATTCACAGATCCACAGCGCAAGAACTGTTCGACAGATGCGCAGCAG 487
QY 919 CTGGGCTCGGCGGAGCTGCGGGCCCGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 978
Db 488 CTGGGCTCGGCGGAGCTGCGGGCCCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 547
QY 979 TCGGGCGCCCGGATGCCAGGCTTGGCTTGCCTTCGCCCGCGCGCGCGCGCGCGCGCG 1038
Db 548 TCGGGCGCCCGGATGCCAGGCTTGGCTTGCCTTCGCCCGCGCGCGCGCGCGCGCGCG 607
QY 1039 GCCTTCGCCAGTGCCTATGCAAGCGGCACGCAAGATGCCAGGCTACGCTGCAGCAAG 1098
Db 608 GCCTTCGCCAGTGCCTATGCAAGCGGCACGCAAGATGCCAGGCTACGCTGCAGCAAG 667
QY 1099 AAGCCCTCGACGTGAACCTCAAGGAGCTGGGCTGGGAGCTGATTAATCGGCCCTTG 1158
Db 668 AAGCCCTCGACGTGAACCTCAAGGAGCTGGGCTGGGAGCTGATTAATCGGCCCTTG 727
QY 1159 GAGTACGAGGCTTATCAGTCGAGGCTGTATCGACTTCCCGCTGCGCTGCACCTGGAG 1218
Db 728 GAGTACGAGGCTTATCAGTCGAGGCTGTATCGACTTCCCGCTGCGCTGCACCTGGAG 787
QY 1219 CCACCAACACGCCATCATCCAGAGCTGTATGAATCCATGAGACCCCGGCTCCACCCCG 1278
Db 788 CCACCAACACGCCATCATCCAGAGCTGTATGAATCCATGAGACCCCGGCTCCACCCCG 847
QY 1279 CCAGTGTGCTGCGGCCACCAAAATGATCCCATCAGCATTCATATACATCAGCGGGC 1338
Db 848 CCAGTGTGCTGCGGCCACCAAAATGATCCCATCAGCATTCATATACATCAGCGGGC 907
QY 1339 AATAATGTGCTTACAGCAGTACGAGGACATGGTGTGAGTCTGCGGCTGCAGGTAG 1398
Db 908 AATAATGTGCTTACAGCAGTACGAGGACATGGTGTGAGTCTGCGGCTGCAGGTAG 967
QY 1399 CGGTGCTTTTCGCCGCGCTTGGCCCG 1425
Db 968 CGGTGCTTTTCGCCGCGCTTGGCCCG 994
```

```
RESULT 5
AAD18333
ID AAD18333 standard; DNA; 1171 BP.
XX
AC AAD18333;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human full length VL-1 or BMP-13 DNA.
KW Human; bone morphogenic protein; BMP-13; vulnery; antiinflammatory;
KW analgesic; gene therapy; transforming growth factor-beta; TGF-beta;
KW tissue formation; wound healing; tissue repair; ligament defect; VL-1;
KW carpal tunnel syndrome; tendonitis; ds.
```

```
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT CDS 2..967
FT /*tag= a
FT /product= "Human VL-1 protein"
FT /note= "CDS does not include start codon"
FT /partial
FT 2..604
FT /*tag= b
FT 605..964
FT /*tag= c
FT /product= "Mature VL-1 protein"
XX
XX US284872-B1.
XX
XX 04-SEP-2001.
XX
XX 28-FEB-1997; 97US-0808324.
XX
XX 22-DEC-1994; 94US-0362670.
PR 07-DEC-1993; 93US-0164103.
PR 02-MAR-1994; 94US-0217780.
PR 02-NOV-1994; 94US-0333576.
XX
XX (GENY ) GENETICS INST INC.
PA (HARD ) HARVARD COLLEGE.
XX
PI Celeste AJ, Wozney JM, Rosen VA, Woliman NM, Thomsen GH;
PI Melton DA;
XX
XX WPI; 2001-588978/66.
DR P-PSDB; AAE10982.
XX
XX New chimeric DNAs, useful for treating tendonitis, carpal tunnel
XX syndrome and other tendon and ligament defects, comprises DNA encoding
XX propeptide linked to DNA encoding bone morphogenetic proteins (BMP)-12,
XX BMP-13 or MP52 -
XX
XX Example 1; Column 55-58; 42pp; English.
XX
XX The invention relates to a chimeric DNA comprising a DNA sequence
XX encoding a propeptide from a member of the transforming growth factor
XX (TGF)-beta superfamily of proteins. This DNA is linked to a DNA sequence
XX encoding an amino acid sequence encoding a mature polypeptide consisting
XX of Bone Morphogenetic Protein (BMP)-12, BMP-13 or MP52 protein. The DNA
XX sequences are useful for producing proteins which induce tendon/ligament
XX like tissue formation, and for isolating and cloning further DNA
XX sequences encoding BMP-12 related proteins with similar activity. The
XX proteins are useful for the induction of tendon/ligament-like tissue
XX formation, wound healing, ligament and other tissue repair, augmenting
XX the activity of bone morphogenetic proteins, and for treating tendonitis,
XX carpal tunnel syndrome and other tendon and ligament defects. The
XX present sequence is human VL-1 DNA fragment also designated as
XX BMP-13.
XX
XX Sequence 1171 BP; 195 A; 384 C; 413 G; 179 T; 0 other;
XX
XX Query Match 69.3%; Score 987; DB 22; Length 1171;
XX Best Local Similarity 100.0%; Pred. No. 1.2e-167;
XX Matches 987; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 439 GATCTCTCGCACACTCTCTCCGGAGACAGAGTATTGTTGATGTGCCATGCTCTCA 498
Db 8 GATCTCTCGCACACTCTCTCCGGAGACAGAGTATTGTTGATGTGCCATGCTCTCA 67
QY 499 GACAAAGAGAGCTGGTGGGCGGAGCTGGGGCTCTTCGCGAGCGCCCTCAGCGCC 558
Db 68 GACAAAGAGAGCTGGTGGGCGGAGCTGGGGCTCTTCGCGAGCGCCCTCAGCGCC 127
QY 559 TGGGGGCCACACAGCGCGCGCTCCACGTGACGTCTTCCCTTGCTTTGCGCCCTACTG 618
|||||
```

Db 128 TGGGGGCCACACCGCGCGCTCCAGTGCAGCTCTTCCCTTGGCTTTCGCCCTTACTG 187  
QY 619 CTGACGCGCGGACCTTGGACCGCGAGGGGCGCCCGCGCGCTGGGAAGTCTTCGAC 678  
Db 188 CTGGACGCGCGGACCTTGGACCGCGAGGGGCGCGCGCTGGGAAGTCTTCGAC 247  
QY 679 GTGTGCGAGGGCTGCGCCACACAGCCCTTGAAGCAGCTGTGCTTGGAGTGGGGCGCA 738  
Db 248 GTGTGCGAGGGCTGCGCCACACAGCCCTTGAAGCAGCTGTGCTTGGAGTGGGGCGCA 307  
QY 739 TGGGCGAGCTGACCGCGGAGGCGGAGCGCGCGCGGAGCCCGAGCAACCGCG 798  
Db 308 TGGGCGAGCTGACCGCGGAGGCGGAGCGCGCGCGGAGCCCGAGCAACCGCG 367  
QY 799 CCCCCGAGCTGCGGAGTCTGGGCTTTCGCGCGGAGGCTGGGCTTCCCGAGGCGGGCC 858  
Db 368 CCCCCGAGCTGCGGAGTCTGGGCTTTCGCGCGGAGGCTGGGCTTCCCGAGGCGGGCC 427  
QY 859 CTGCTGGTGTATTCACAGATCCCGAGCGGAGCAAGCTGTTCGACAGATGGCGGAG 918  
Db 428 CTGCTGGTGTATTCACAGATCCCGAGCGGAGCAAGCTGTTCGACAGATGGCGGAG 487  
QY 919 CTGGGCTGCGCGAGCTGCGGCGCGGCGCGGCGCGGCGCGGCGCGGCGCG 978  
Db 488 CTGGGCTGCGCGAGCTGCGGCGCGGCGCGGCGCGGCGCGGCGCGGCGCG 547  
QY 979 TCGGGCGCGCGGATGCCAGGCTTGGCTTCCCTCGCGCGCGCGCGCGCGCG 1038  
Db 548 TCGGGCGCGCGGATGCCAGGCTTGGCTTCCCTCGCGCGCGCGCGCGCGCG 607  
QY 1039 GCCTTCGCGAGTGCATGCGAAGCGCGCGCGCGCAAGTCCAGGCTACGCTGCAGCA 1098  
Db 608 GCCTTCGCGAGTGCATGCGAAGCGCGCGCGCGCAAGTCCAGGCTACGCTGCAGCA 667  
QY 1099 AAGCCCTGCAGTGAACCTCAAGGAGCTGGCTGGGAGCTGGATTATCGGCCCTTG 1158  
Db 668 AAGCCCTGCAGTGAACCTCAAGGAGCTGGCTGGGAGCTGGATTATCGGCCCTTG 727  
QY 1159 GAGTAGAGGCTTATCAGTCCGAGGCTGTATGCGACTTCCCGTCCGACTCGGAG 1218  
Db 728 GAGTAGAGGCTTATCAGTCCGAGGCTGTATGCGACTTCCCGTCCGACTCGGAG 787  
QY 1219 CCACCAACGCGCATATCAGAGCGCTGATGAATCCATGACCCCGGCTCCACCCCG 1278  
Db 788 CCACCAACGCGCATATCAGAGCGCTGATGAATCCATGACCCCGGCTCCACCCCG 847  
QY 1279 CCAGCTGTGCTGCGCCCAACCAATTTGACTCCCATCAGCATCTATACATCGAGCG 1338  
Db 848 CCAGCTGTGCTGCGCCCAACCAATTTGACTCCCATCAGCATCTATACATCGAGCG 907  
QY 1339 AATAATGTGCTTACAGCAGTACGAGGACATGGTGTGGAGTCTGCGGCTGCAGGTAG 1398  
Db 908 AATAATGTGCTTACAGCAGTACGAGGACATGGTGTGGAGTCTGCGGCTGCAGGTAG 967  
QY 1399 CGGTGCTTTCGCGCGCGCTTGGCCCG 1425  
Db 968 CGGTGCTTTCGCGCGCGCTTGGCCCG 994

RESULT 6

AAQ96208

ID AAQ96208 standard; DNA; 1171 BP.

XX

AC

XX

DT

XX

DE

XX

KW

XX

OS

XX

AAQ96208;

30-NOV-1995 (first entry)

Human mature VL-1 (BMP-13) encoding sequence.

Bone morphogenetic protein; VL-1; tendon; ligament; tendinitis; ss.

Homo sapiens.

Key Location/Qualifiers

FT CDS 2..964  
FT /tag= a  
FT mat\_peptide 605..964  
FT /tag= b  
FT misc\_difference 671..874  
FT /tag= c  
FT /note= "degenerate codes used due to poor print quality"  
FT misc\_feature 605..964  
FT /tag= d  
FT /note= "claimed"  
FT misc\_feature 659..964  
FT /tag= e  
FT /note= "claimed"  
XX W09516035-A.  
XX 15-JUN-1995.  
XX 06-DEC-1994; 94WO-US14030.  
XX 02-NOV-1994; 94US-0333576.  
XX 07-DEC-1993; 93US-0164103.  
XX 25-MAR-1994; 94US-0217780.  
XX (GEM ) GENETICS INST INC.  
XX (HARD ) HARVARD COLLEGE.  
XX Celeste AJ, Melton DA, Rosen VA, Thomsen GH, Wolfman NM;  
XX Wozney JM;  
XX P-PSDB; AAR78730.  
XX WPI: 1995-224320/29.  
XX Bone morphogenetic proteins -12 and -13 and corresp. DNA - used in  
XX compsn. for inducing tendon/ligament-like tissue formation  
XX Claim 2; Page 62-64; 84pp; English.  
XX BMP-12 related proteins are a subset of the BMP/TGF-beta/Vg-1  
XX family of proteins, including BMP-12 and VL-1. VL-1 is designated  
XX BMP-13. Like BMP-12, it is expected that BMP-13, as expressed by  
XX mammalian cells such as CHO cells, exists as a heterogeneous popn.  
XX of active species of BMP-13 protein with varying N-termini. It is  
XX expected that all active species will contain the AA sequence  
XX beginning with the 19th Cys residue of the mature protein until  
XX the 119th residue of the mature protein or until the stop codon  
XX after the 120th residue of the mature protein. Other active  
XX species contain additional AA sequence in the N-terminal direction.  
XX CC AAQ96208 is a partial DNA sequence and AAR78730 is the derived AA  
XX sequence of a portion of the 2.5 kb DNA insert of the plasmid  
XX subclone pGEMJIDc31/2.5, derived from clone lambdaJIDc31.  
SQ Sequence 1171 BP; 195 A; 377 C; 408 G; 174 T; 4 U; 13 other;

Query Match

Best Local Similarity 68.5%; Score 976; DB 16; Length 1171;

Matches 969; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 439 GATCTCTCGCACACTCTCTCCGGAGACAGAAAGTATTTTGTGATGTGCTTCA 498  
Db 8 GATCTCTCGCACACTCTCTCCGGAGACAGAAAGTATTTTGTGATGTGCTTCA 67  
QY 499 GACAAAGAGAGCTGTGGGCGGAGCTTTCGGCAGCGCCCTCAGGCGCC 558  
Db 68 GACAAAGAGAGCTGTGGGCGGAGCTTTCGGCAGCGCCCTCAGGCGCC 127  
QY 559 TGGGGGCCACCGCGCGGCTCCACGTGCAGCTTCCCTTGCCTTTCGCCCTTACTG 618  
Db 128 TGGGGGCCACCGCGCGGCTCCACGTGCAGCTTCCCTTGCCTTTCGCCCTTACTG 187  
QY 619 CTGGACGCGCGGACCTGGACCGCGCGCGCGCGCTGGGAAGTCTTCGAC 678

```
Db 188 CTGGACGCGGACCCCTGGACCGCAGGGGGCGCCCGCGCGCTGGGAAGTCTTCGAC 247
Qy 679 GTGTGGCAGGGCTCGCGCCACAGCCTCGAAGCAGCTGTGCTGGAGCTGGGGCGCA 738
Db 248 GTGTGGCAGGGCTCGCGCCACAGCCTCGAAGCAGCTGTGCTGGAGCTGGGGCGCA 307
Qy 739 TGGGGGAGCTGGACCGCGGGAGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCG 798
Db 308 TGGGGGAGCTGGACCGCGGGAGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCG 367
Qy 799 CCCCCGAGCTCGGGAGCTGGGCTTTCGCGGAGGCTCGGCGCTCCCGAGGCGGGCC 858
Db 368 CCCCCGAGCTCGGGAGCTGGGCTTTCGCGGAGGCTCGGCGCTCCCGAGGCGGGCC 427
Qy 859 CTGCTGGTGGTATTCACAGATCCACAGCAGCAAGAACCTGTCGACAGATGCGGAGCAG 918
Db 428 CTGCTGGTGGTATTCACAGATCCACAGCAGCAAGAACCTGTCGACAGATGCGGAGCAG 487
Qy 919 CTGGGCTCGCGGAGGCTCGCGGCGCGGCGCGGCGGCGGCGGCGGCGGCGGCGG 978
Db 488 CTGGGCTCGCGGAGGCTCGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 547
Qy 979 TCGGGCGCGCGGATGCCAGGCTTGGCTGCGCTCGCGGCGGCGGCGGCGGCGGCGG 1038
Db 548 TCGGGCGCGCGGATGCCAGGCTTGGCTGCGCTCGCGGCGGCGGCGGCGGCGGCGG 607
Qy 1039 GCTTCGCGAGTCGCATGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1098
Db 608 GCTTCGCGAGTCGCATGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 667
Qy 1099 AAGCCCTCGAGTCAAGGAGCTGGGCTGGGAGCTGGATATTCGCGGCGGCGGCGG 1158
Db 668 AAGCCCTCGAGTCAAGGAGCTGGGCTGGGAGCTGGATATTCGCGGCGGCGGCGG 727
Qy 1159 GAGTAGGCGCTATCAGTGGAGGCTGTATGGAGTTCGCGGCTGGGCTGGGCTGGGAG 1218
Db 728 GAGTAGGCGCTATCAGTGGAGGCTGTATGGAGTTCGCGGCTGGGCTGGGCTGGGAG 787
Qy 1219 CCACCAACGCGCATATCAGAGCTGTATGAGTCCATGAGCCCGCGGCTCCACCGCG 1278
Db 788 CCACCAACGCGCATATCAGAGCTGTATGAGTCCATGAGCCCGCGGCTCCACCGCG 847
Qy 1279 CCAGTGTGCTGCGCGCCACCAATGACTCCATCAGCATTCATATACATCAGCGCGGCG 1338
Db 848 CCAGTGTGCTGCGCGCYACNAARUNACTCCATCAGCATTCATATACATCAGCGCGGCG 907
Qy 1339 AATAATGTGGTCTACAGCAGTACGAGGACATGGTGGAGTGTGGGCTGCGGCTGCGAGGTAG 1398
Db 908 AATAATGTGGTCTACAGCAGTACGAGGACATGGTGGAGTGTGGGCTGCGGCTGCGAGGTAG 967
Qy 1399 CGGTGCTTTCCGCGCGGCTTGGCGCG 1425
Db 968 CGGTGCTTTCCGCGCGGCTTGGCGCG 994
```

## RESULT 7

AAT31602

ID AAT31602 standard; DNA; 1308 BP.

XX AAT31602;

AC AAT31602;

XX AAT31602;

DT 25-OCT-1996 (first entry)

XX Cartilage-derived morphogenetic protein-2 gene.

DE Cattle; cartilage-derived morphogenetic protein-2; CDMP-2;

XX articular cartilage; chondrogenic; vulnery; implantation;

KW chondromalacia; osteoarthritis; therapy; joint repair; ss.

XX Bos taurus.

OS Bos taurus.

XX Location/Qualifiers

FH Key

```
FT mat_peptide 949..1308
FT /*tag= a
FT /product= Mature CDMP-2 C-terminal domain
FT misc_feature 1042..1146
FT /*tag= b
FT /note= "RT-PCR-amplified probe used to screen cDNA
FT library"
XX WO9614335-A1.
XX PN 17-MAY-1996.
XX PD 07-NOV-1994; 94WO-US12814.
XX PF 07-NOV-1994; 94WO-US12814.
XX PR 07-NOV-1994; 94WO-US12814.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Chang SC, Luyten FP, Moos M;
XX DR WPI; 1996-251714/25.
XX DR P-PSDB; AAR95636.
XX PT New purified cartilage extracts and proteins - used to stimulate the
XX development and repair of cartilage in vivo.
XX PS Claim 10; Fig 2; 34pp; English.
XX CC The sequence encodes cattle articular cartilage-derived morphogenetic
XX protein-2 (CDMP-2). Primers AAT13603-06, constructed from conserved
XX motifs from bone morphogenetic protein, have been used in reverse
XX CC transcription-polymerase chain reaction amplification of a cattle
XX CC articular chondrocyte cDNA library, generating products of 120 and
XX CC 280 bp, which are then used to screen a cattle articular cartilage
XX CC cDNA library. The 5'-end with the N-terminal methionine codon and
XX CC signal peptide sequence is missing. CDMP-2 is present in a purified
XX CC cartilage extract (claimed) which stimulates local cartilage formation
XX CC and repair when combined with a matrix and implanted in a mammal. The
XX CC protein may be used in therapy of e.g. chondromalacia or
XX CC osteoarthritis, to heal joint surfaces, or to repair cartilage after
XX CC reconstructive surgery.
XX SQ Sequence 1308 BP; 209 A; 456 C; 452 G; 191 T; 0 other;
Query Match 65.1%; Score 928.2; DB 17; Length 1308;
Best Local Similarity 86.5%; Pred. No. 3.8e-157;
Matches 1136; Conservative 0; Mismatches 123; Indels 54; Gaps 8;
Qy 128 COTCCGCGGAGCTGGTTCACCAAGGGCATGCGAAGCGCGCAAGGAAGCAAGATGCGAGC 187
Db 5 COTCCGCGGAGCTGGGCTCCGCGCAAGGGAATGCGAAGCGCGCAAGGAAGGATGCGCGC 64
Qy 188 GGGCGCGCGGAGACAGTACGCGGGCGCGGGAG-----GGCCAGGAACACAGCGCGC 238
Db 65 GGGCGCGGAGAGAGATGCCAGCGCGCGGAGCCCTGGATGCGCGAGGAGCGCGCGCGCA 124
Qy 239 GGCCTCAGAGCAAGAACCC-----CGGGCTCAGCAGCCCGCGCGGCGGAGGCGCGCAG 289
Db 125 GGCCTCAGAGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 184
Qy 290 GCAGGGGTCGCGCGCTGGTGGCGCCACAGTAGTACATGCTGTCAATCTACAGGACTTACTCCA 349
Db 185 GCAGGGGTCGCGCGCTGGTGGCGCCACAGTAGTACATGCTGTCAATCTACAGGACTTACTCCA 244
Qy 350 TCGCTGACAAGCTGGGCATCAATGCCAGCTTTTCCAGTCTTCCAGTCCGCTAATAGA 409
Db 245 TCGCGGAGAGCTGGGCATCAATGCTAGCTTTTCCAGTCTTCCAGTCCGCTAATAGA 304
Qy 410 TCACAGCTTTGTAGACAGGGGACTAGACGATCTCTCCACACTCTCTCCGAGACAGAGA 469
Db 305 TCACAGCTTTGTAGACAGGGGACTAGACGATCTCTCCACACTCTCTCCGAGACAGAGA 364
Qy 470 AGTATTTGTTGATGTGTCATGCTCTCAGACAAAGAGAGCTGTGGGCGCGGAGCTGC 529
```





Db 347 GGAGCCGAGGGGTCAATGTCACGCGCGTGGGGTCCGCCAGACACCGGGTCTTGGCTGCC 406

QY 1012 TCGCCGGCGCGCGGGGGCGACGGCTTCGCCAGTCGCCATGCAAGCGGCACGGC 1071

Db 407 TCGCCGGCGCGCGGGGGCGACCGCTTCGCCAGCGGTACCGCAAGCAGCATGGC 466

QY 1072 AAGAAGTCCAGGCTAGCGTGCAGCAAGAACGCCCTCGACGTGAACCTCAAGGAGCTGGGC 1131

Db 467 AAGAAGTCCAGGCTAGCGTGCAGCAAGAACGCCCTCGACGTGAATTTAAGGAGTTAGGC 526

QY 1132 TGGGAGACTGGATTATCGCGCCCTTGGAGTAGAGAGCCCTATCACTGCGAGGGTGTATGC 1191

Db 527 TGGGAGACTGGATTATCGCGCCCTTAGAGTAGAGAGCCCTATCACTGCGAGGGCGTGTGC 586

QY 1192 GACTTCCCGCTGGCTCGACCTGGAGCCACCAACAGCCCATCATCAGAGCGTGTATG 1251

Db 587 GACTTCCCGCTGGCTCGACCTGGAGCCACCAACAGCCCATCATCAGAGCGTGTATG 646

QY 1252 AACTCCATGAGCCCGGCTCCACCCCGCCAGCTGCTGGTGCCCAACCAATTGACTCCC 1311

Db 647 AACTCCATGAGCCCGGCTCCACCCCGCCAGCTGCTGGTGCCCAACCAACTGACTCCC 706

QY 1312 ATCAGCATCTATACATCAGCGGGGCAATATGTGTTACAGCAGTACAGGACATG 1371

Db 707 ATTAGCATCTGTACATCAGCGGGGCAATATGTGTTACAGCAGTATGAGGACATG 766

QY 1372 GTGGTGGAGTCGCGGCTGCAGGTAGCGTCTTCCGCGCGCTTGGCCCG 1425

Db 767 GTGGTGGAGTCGCGGCTGTAGTAGCGGTG-CTGTCCCGCCACCTGGGCCAG 819

RESULT 9

AAD18336

ID AAD18336 standard; DNA; 1046 BP.

XX AC AAD18336;

XX DT 18-DEC-2001 (first entry)

XX DE Murine partial mv2 DNA.

XX KW Mouse; bone morphogenic protein; BMP-12; vulnery; antiinflammatory; analgesic; gene therapy; transforming growth factor-beta; TGF-beta;

XX KW tissue formation; wound healing; tissue repair; ligament defect;

XX KW carpal tunnel syndrome; tendonitis; mv2; ds.

OS Mus sp.

XX FH Key

XX FT 2..793

XX FT CDS

XX FT Location/Qualifiers

XX FT /\*tag= a

XX FT /product= "Murine partial mv2 protein"

XX FT /transl\_except= (pos:160..163, aa:Xaa)

XX FT /transl\_except= (pos:740..742, aa:Xaa)

XX FT /note= "Xaa is an unknown amino acid; CDS does not include start codon"

XX FT /partial

XX

PN US6284872-B1.

XX

PD 04-SEP-2001.

XX

XX 28-FEB-1997; 97US-0808324.

XX

XX 22-DEC-1994; 94US-0362670.

XX 07-DEC-1993; 93US-0164103.

XX 25-MAR-1994; 94US-0217780.

XX 02-NOV-1994; 94US-0333576.

XX

PA (GENY ) GENETICS INST INC.

PA (HARD ) HARVARD COLLEGE.

XX Celeste AJ, Wozney JM, Rosen VA, Wolfman NM, Thomsen GH;

PI Melton DA;

XX

DR WPI: 2001-588978/66.

DR P-PSDB; AAE10985.

XX

PT New chimeric DNAs, useful for treating tendonitis, carpal tunnel

PT syndrome and other tendon and ligament defects, comprises DNA encoding

PT propeptide linked to DNA encoding bone morphogenetic proteins (BMP)-12,

PT BMP-13 or MP52 -

XX

PS Example 1; Column 71-74; 42pp; English.

XX

CC The invention relates to a chimeric DNA comprising a DNA sequence

CC encoding a propeptide from a member of the transforming growth factor

CC (TGF)-beta superfamily of proteins. This DNA is linked to a DNA sequence

CC encoding an amino acid sequence encoding a mature polypeptide consisting

CC of Bone Morphogenetic Protein (BMP)-12, BMP-13 or MP52 protein. The DNA

CC sequences are useful for producing proteins which induce tendon/ligament

CC like tissue formation, and for isolating and cloning further DNA

CC sequences encoding BMP-12 related proteins with similar activity. The

CC proteins are useful for the induction of tendon/ligament-like tissue

CC formation, wound healing, ligament and other tissue repair, augmenting

CC the activity of bone morphogenetic proteins, and for treating tendonitis,

CC carpal tunnel syndrome and other tendon and ligament defects. The

CC present sequence is murine partial mv2 DNA which is homologous to

CC human BMP-12 or VL-1 sequences of the invention.

XX

SQ Sequence 1046 BP; 212 A; 314 C; 331 G; 187 T; 2 other;

Query Match 40.8%; Score 581.8; DB 22; Length 1046;

Best Local Similarity 87.2%; Pred. No. 2.9e-95;

Matches 675; Conservative 1; Mismatches 84; Indels 14; Gaps 3;

QY 653 CGCGCGCGCGCTGGGAAGTC-TTCGACGTGTGCGAGGGCTGGCCACAGCCCTGGAAG 711

Db 59 CCCAGCGCGCTGGGAAGTC-TTCGACGTGTGCGAGGGCTGGCCCTCAGCCTTGGGAAG 118

QY 712 CAGCTGTGCTGTGAGCTGGCGCGCATGCGGCGAGCTGCGAGCGGAGGCGGAGCGG 771

Db 119 CAGCTGTGCTGTGAGCTGGCGCGCATGCGGCGAGCTGGGTTGACACCGGGGATAGCGGGCG 178

QY 772 CGCGCGCGGGGACCCAGCAACCGCGCGCGGAGCTCGGAGTCTGGGCTTCGCCCGG 831

Db 179 CGCGCGAGGGGTCCCGCAGCAGCACCGCCTCTGGACCTCGGAGTCTGGCTTCGCTCG 238

QY 832 AGGTGCGGCTTCCCGCAGGAGGGCCCTGCTGGTGGTATTACACAGATCCAGCGCCAG 891

Db 239 AGGTGAGACCGCGCGCGAGGCGCGCTGTGTGTGTTCACCAAGATCGCAGCGCAAG 298

QY 892 AACCTGTTCCGAGATGCGCGAGCTGGGCTCGGCGGAGGCTCGGCGCGCGCGCGGCG 951

Db 299 AACCTGTTCCGAGATGATGATGAGCAGCTGGGCTCTGCAGAGGCT-----GCG 346

QY 952 GCGCGCGAGGGTCTGTGGCGCGCGCGCTCGGCGCGCGCGGATGCCAGGCTTGGCTGCC 1011

Db 347 GGAGCGAGGGGTCTATGTCACGCGCGTGGGCTCCCGCAGACACCGGGTCTTGGCTGCC 406

QY 1012 TCGCGCGCGCGCGGGGCGGCGAGCGGCTTCGCCAGTTCGCCATGCGGCGGCGGCGGCG 1071

Db 407 TCGCGCGCGCGCGGGGCGGCGAGCGCGCTTCGCCAGCGCTTCGCCAGGACATGCGC 466

QY 1072 AAGAAGTCCAGGCTAGCTGCGAGCAAGAACGCCCTGCGAGTGAACCTCAAGGAGCTGGGC 1131

Db 467 AAGAAGTCCAGGCTAGCTGCGAGCAAGAACGCCCTGCGAGTGAATTTAAGGAGTTAGGC 526

QY 1132 TGGGAGCACTGGATTATCGCGCCCTTGGAGTAGAGGCGCTATCACTGCGAGGGTGTATGC 1191

Db 527 TGGGAGCACTGGATTATCGCGCCCTTAGAGTAGAGGCGCTATCACTGCGAGGGCGTGTGC 586

QY 1192 GACTTCCCGCTGGCTCGACCTGGAGCCACCAACAGCCCATCATCAGAGCGTGTATG 1251

Db 587 GACTTCCCGCTGGCTCGACCTGGAGCCACCAACAGCCCATCATCAGAGCGTGTATG 646

QY 1252 AACTCCATGACCCCGCTCCACCCCGCCAGCTGCTGCGTGCCACCAAAATTGACTCCC 1311  
Db 647 AACTCCATGACCCCGCTCCACCCCGCCAGCTGCTGCGTGCCACCAAAATTGACTCCC 706

QY 1312 ATCAGCATTCATACATCGACGGGGCAATAATGTGCTTACAGCAGTACGAGGACATG 1371  
Db 707 ATTAGCATCTGTACATCGACGGGGCAATAATGTGCTTACAGCAGTACGAGGACATG 766

QY 1372 GTGGTGAGTCTGCGGCTGCAGCTAGCTGCTTTCGCCCGCCCTTGCCCGG 1425  
Db 767 GTGGTGAGTCTGCGGCTGTAGGTAGCGGTG-CTGTCCCGCCACTTGGGCCAG 819

RESULT 10  
AAQ96224  
ID AAQ96224 standard; DNA; 1046 BP.  
XX  
AC AAQ96224;  
XX  
DT 23-NOV-1995 (first entry)  
DE Murine mv2 protein.  
XX  
KW Bone morphogenetic protein; mv2; tendon; ligament; ss.  
XX  
OS Mus musculus.  
XX  
FH Location/Qualifiers  
FT CDS 2..790  
FT misc\_difference 161 /\*tag= a  
FT /\*tag= b  
FT /\*label= C or A  
FT misc\_difference 742 /\*tag= c  
FT /\*label= A,C,T or G  
XX  
XX W09516035-A.  
XX  
XX 15-JUN-1995.  
XX  
XX 06-DEC-1994; 94WO-US14030.  
XX  
XX 02-NOV-1994; 94US-0333576.  
XX 07-DEC-1993; 93US-0164103.  
XX 25-MAR-1994; 94US-0217780.  
XX  
XX (GENY ) GENETICS INST INC.  
XX (HARD ) HARVARD COLLEGE.  
XX  
XX Celeste AJ, Melton DA, Rosen VA, Thomsen GH, Wolfman NM;  
XX Wozney JM;  
XX  
XX WPI; 1995-224320/29.  
XX P-PSDB; AAR78739.  
XX  
XX Bone morphogenetic proteins -12 and -13 and corresp. DNA - used in  
XX compsn. for inducing tendon/ligament-like tissue formation  
XX  
XX Example; page 71-72; 84pp; English.  
XX  
XX Oligos #6 and #7 (AAQ96218 & AAQ96219) are used as primers for the  
XX amplification of a 275 bp DNA probe, the internal 269 bp of which  
XX corresp. to nts #607 to #865 of AAQ96207, from the BMP-12 encoding  
XX plasmid subclone PCR1-l#2. This probe was radioactively labelled  
XX and used to screen a murine genomic library. DNA sequence analysis  
XX of one of positively hybridising recombinants named MVR32 indicates  
XX that it encodes a portion of the mouse gene corresp. to the PCR  
XX product mv2 (murine homolog of the human VL-1 sequence AAQ96213.  
XX The partial DNA sequence of this subclone and corresp. AA  
XX translation are given in AAQ96224 & AAR78739.

Sequence 1046 BP; 212 A; 314 C; 331 G; 187 T; 2 other;

Query Match 40.8%; Score 581.2; DB 16; Length 1046;  
Best Local Similarity 87.2%; Pred. No. 3.7e-95;  
Matches 675; Conservative 0; Mismatches 85; Indels 14; Gaps 3;

QY 653 CGCGCGCGGCTGGGAAGTC-TTGCAGCTGTGGCAGGCGCTGCGCCACAGCCCTGGAAG 711  
Db 59 CCCAGCGCGCTGGGAAGTCTTTTCGACGTGTGGCAGGCGCTGCGCCCTGAGCCCTTGAAG 118

QY 712 CAGCTGTGCTTGGAGCTGCGGGCCGATGGGCGAGCTGGACGCCGGGGGCGGAGCG 771  
Db 119 CAGCTGTGCTTGGAGTTGCGGGCAGCCTGGGTGAGTGACNCCGGGGATACGGGGCG 178

QY 772 CGCGCGGGGACCCACAGAACCGCCCGCCGACCTGCGGAGTCTGGGCTTCGGCGG 831  
Db 179 CGCGGAGGGGTCCCAAGAGCCACCGCCTTGACCTCGGAGTCTGGGCTTCGGTCG 238

QY 832 AGGTGCGGCTCCCGAGGAGCGGCGCTGTGTGTATTCACAGATCCACGCGCAAG 891  
Db 239 AGGTGAGACCGCCCGAGGAGCGCGCTGTGTGTATTCACAGATCCACGCGCAAG 298

QY 892 AACCTGTTCCAGAGATGCGGAGCAGCTGGGCTCGGCCGAGGCTGCGGGCCGGCGG 951  
Db 299 AACCTGTTCACTGAGATGCATGACGTGGGCTCTGCAGAGGCT-----GCG 346

QY 952 GCGCGCAGGGGTGCTGCGCGCGCGCCCTCGGGCGCCCGGATGCCAGGCTTTGGCTGCC 1011  
Db 347 GGAGCCGAGGGTTCATGTCCAGCGCGCTCGGGCTCCCGACACCGGCTTTGGCTGCC 406

QY 1012 TCGCGCGCGCGCGCGCGCGCGCGCGCTTCGCCAGTGCCTATGCAAGCGCAGCGG 1071  
Db 407 TCGCGCGCGCGCGCGCGCGCGCGCTTCGCCAGCGCTCACGGAAGCGACATGGC 466

QY 1072 AAGAGTCCAGGCTACGCTGCAGCAAGAGCCCTGCACGTGAATTCAGGAGCTGGG 1131  
Db 467 AAGAGTCCAGGCTGCGCTGCAGCAAGAGCCCTGCACGTGAATTTAAGGAGTTAGG 526

QY 1132 TGGGACGACTGGATTATCGCGCGCGCTGGAGTACGAGGCTTATCACTCGAGGGTGTATG 1191  
Db 527 TGGGACGACTGGATTATCGCGCGCGCTAGAGTACGAGGCTTATCACTCGAGGGGTGTC 586

QY 1192 GACTTCCCGCTGCGCTCGCACCTGGAGCCCAACACGCGCATATCCAGAGCGCTGATG 1251  
Db 587 GACTTCCCGCTGCGCTCGCACCTGGAGCCCACTAACCATGCCATCATTCAGACGCTGATG 646

QY 1252 AACTCCATGGACCCCGCTCCACCCCGCCAGCTGCTGCGTGGCCCAAAATTGACTCCC 1311  
Db 647 AACTCCATGGACCCCGCTCCACCCCGCCAGTGTGCTGCGTTCCCAAACTGACTCCC 706

QY 1312 ATCAGCATTCATACATCGACGGGGCAATAATGTGCTTACAAAGCAGTACGAGGACATG 1371  
Db 707 ATTAGCATCTGTACATCGACGGGGCAATAATGTGCTTACAAAGCAGTACGAGGACATG 766

QY 1372 GTGGTGAGTCTGCGGCTGCAGGTAGCGGTGCGCTTTCGCCCGCCCTTGGGCCG 1425  
Db 767 GTGGTGAGTCTGCGGCTGTAGGTAGCGGTG-CTGTCCCGCCACTTGGGCCAG 819

RESULT 11  
AAQ84280  
ID AAQ84280 standard; DNA; 530 BP.  
XX  
AC AAQ84280;  
XX  
DT 17-AUG-1995 (first entry)  
DE  
XX GDF-6 gene.  
XX  
XX Growth differentiation factor-6; GDF-6;  
XX transforming growth factor-beta; TGF-beta; cell proliferation;  
XX pancreas; ss.  
XX  
XX Mus sp.

```
XX FH Key Location/Qualifiers
XX FT CDS 126..530
XX FT /*tag= a
XX
XX PN W09501801-A.
XX
XX PD 19-JAN-1995.
XX
XX PF 08-JUL-1994; 94WO-US07762.
XX
XX PR 09-JUL-1993; 93US-0089300.
XX
XX PA (UYJO ) UNIV JOHNS HOPKINS SCHOOL MED.
XX
XX PI Huynh T, Lee S;
XX
XX DR WPI; 1995-066738/09.
XX DR P-PSDB; AAR66867.
XX
XX PT Growth differentiation factor-6 - a member of the TGF-beta
XX PT family, useful to treat cell proliferation disorders in
XX PT pancreatic tissue
XX
XX PS Disclosure; Page 31-32; 62pp; English.
XX
XX CC Primer SJL141 (given in AAQ84278), based on a region spanning 2 Trp
XX CC residues conserved in TGF-beta proteins, and primer SJL145 (AAQ84279)
XX CC spanning the invariant Cys residues near the C-terminus were used
XX CC to amplify mouse genomic DNA. New growth differentiation factor-6
XX CC (AAQ84280) was identified from a mixture of PCR products.
XX
XX SQ Sequence 530 BP; 104 A; 166 C; 160 G; 100 T; 0 other;

Query Match 28.9%; Score 412.2; DB 16; Length 530;
Best Local Similarity 87.8%; Pred. No. 5.8e-65;
Matches 477; Conservative 0; Mismatches 53; Indels 13; Gaps 2;

QY 856 GCCTGTGTGGTATTACACAGATCCAGCGCAAGAACCTGTCGAGAGATGCCGAG 915
DB 1 GGCCTGTGTAGTGTTCACAGATCCAGCGCAAGAACCTGTTCACTGAGATGCATGAG 60

QY 916 CAGCTGGGCTCGGCCAGGCTCGGGCCCGGGCCCGGGGTCGTGGCCGCCG 975
DB 61 CAGCTGGGCTCTCAGAGGCT-----GCGGGAGCCGAGGGGTCATGGCCAGCG 108

QY 976 CGCTCGGGCCCGCGGATCCAGGCTTGGCTCGCCCTCGCCGCGCGCGCGCGCGC 1035
DB 109 CCGTCCGGGC-TCCACAGCCCGGGTCTTGGCTGCCCTCGCCCGCGCGCGCGCACGC 167

QY 1036 ACGGCCCTTCGCCAGTCCGATGCCAAGCGCACGCAAGATCCAGGCTACGCTGCAGC 1095
DB 168 ACCGCCCTTCGCCAGCGCTCAGCGACATGGCAAGATCCAGGCTGCAGC 227

QY 1096 AAGAAGCCCTTCAGCTGACATTCAGGAGCTGGGCTGGGACGACATGATATCGGCC 1155
DB 228 AAGAAGCCCTTCAGCTGACATTCAGGAGCTGGGCTGGGACGACATGATATCGGCC 287

QY 1156 CTGGAGTACGAGGCTTACCTCGAGGGTGTATGCGACTTCCGCTCGGCTCGCACCTG 1215
DB 288 CTAGAGTACGAGGCTTACCTCGAGGGGCTGTGCGACTTTCCGCTCGGCTCGCACCTT 347

QY 1216 GAGCCACCAACACGCCATCATCCAGAGCGCTGATGAATCCATGACCCCGGCTCCACC 1275
DB 348 GAGCCACCAATCAATCCATTCATCAGAGCGCTGATGAATCCATGACCCCGGCTCCACC 407

QY 1276 CCGCCAGCTGCTGCGTCCGCCAATATGACTTCCCATCAGCATTTCTATACATCGACGG 1335
DB 408 CCGCTAGCTGCTGCTTCCGCCAATATGACTTCCCATCAGCATTTCTATACATCGACGG 467

QY 1336 GGCATTAATGTGCTTACAGAGATACGAGGACATGGTGGTGGAGTCTGCGGCTCAGG 1395
DB 468 GGCATTAATGTGCTTACAGAGATACGAGGACATGGTGGTGGAGTCTGCGGCTCAGG 527
```

```
QY 1396 TAG 1398
DB 528 TAG 530

RESULT 12
AAA72621
ID AAA72621 standard; DNA; 530 BP.
XX
XX AC AAA72621;
XX
XX DT 29-NOV-2000 (first entry)
XX
XX DE Murine growth differentiation factor-6 (GDF-6) DNA sequence.
XX
XX KW Antibody; growth differentiation factor-6; GDF-6; placenta; mouse;
XX KW transforming growth factor-beta; TGF-beta; choriocarcinoma; neoplasm;
XX KW detection; treatment; cell proliferative disease; ds.
XX
XX OS Mus sp.
XX
XX PN US6090563-A.
XX
XX PD 18-JUL-2000.
XX
XX PF 15-JUN-1998; 98US-0097616.
XX
XX PR 15-APR-1996; 96US-0581529.
XX
XX PA (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
XX PI Lee S, Huynh T;
XX
XX DR WPI; 2000-531418/48.
XX DR P-PSDB; AAB12986.
XX
XX PT Antibody for detecting cell proliferative disorder associated with
XX PT growth differentiation factor-6 expression, specifically binds to
XX PT growth differentiation factor-6 polypeptide
XX
XX PS Example 2; Fig 2; 24pp; English.
XX
XX CC This invention relates to an antibody that specifically binds to a growth
XX CC differentiation factor-6 (GDF-6) polypeptide. GDF-6 is a member of the
XX CC transforming growth factor-beta (TGF-beta) family of proteins, and is
XX CC expressed in placental tissue. The antibody of the invention can be used
XX CC to identify tissue samples of placental origin. It may also be used to
XX CC treat choriocarcinomas, or detect primary and metastatic neoplasms of
XX CC placental origin. The antibodies may also be used as an indicator of
XX CC proliferative diseases, and GDF-6 may be used as an indicator of
XX CC developmental anomalies in prenatal screening procedures. The present
XX CC sequence represents DNA encoding the murine GDF-6 protein.
XX
XX SQ Sequence 530 BP; 104 A; 166 C; 160 G; 100 T; 0 other;
```

```
Query Match 28.9%; Score 412.2; DB 21; Length 530;
Best Local Similarity 87.8%; Pred. No. 5.8e-65;
Matches 477; Conservative 0; Mismatches 53; Indels 13; Gaps 2;

QY 856 GCCTGTGTGGTATTACACAGATCCAGCGCAAGAACCTGTCGAGAGATGCCGAG 915
DB 1 GGCCTGTGTAGTGTTCACAGATCCAGCGCAAGAACCTGTTCACTGAGATGCATGAG 60

QY 916 CAGCTGGGCTCGGCCAGGCTCGGGCCCGGGCCCGGGGTCGTGGCCGCCG 975
DB 61 CAGCTGGGCTCTCAGAGGCT-----GCGGGAGCCGAGGGGTCATGGCCAGCG 108

QY 976 CGCTCGGGCCCGCGGATCCAGGCTTGGCTCGCCCTCGCCGCGCGCGCGCGCGC 1035
DB 109 CCGTCCGGGC-TCCACAGCCCGGGTCTTGGCTGCCCTCGCCCGCGCGCGCGCACGC 167

QY 1036 ACGGCCCTTCGCCAGTCCGATGCCAAGCGCACGCAAGATCCAGGCTACGCTGCAGC 1095
```

```
Db 168 ACCGCTTCCCGACCGCTACGCGCAAGCGACATGGCAAGATCCAGGCTGCCTGCAGC 227
QY 1096 AAGAAGCCCTCGACGCTGAACCTCAAGGAGCTGGCTGGGACGACTGGATTATCGCGGCC 1155
Db 228 AGAAGCCCTCGACGCTGAATTTTAAGGAGTTAGGCTGGGACGACTGGATTATCGCGGCC 287
QY 1156 CTGAGTACGAGGCTTATCACTACGAGGCTGTATGGGACTTCCGCTGCGCTGCACCTG 1215
Db 288 CTAGAGTACGAGGCTTATCACTACGAGGCTGTATGGGACTTCCGCTGCGCTGCACCTT 347
QY 1216 GAGCCCAACCAACGACATCACTACGAGGCTGTATGGGACTTCCGCTGCGCTGCACCT 1275
Db 348 GAGCCCAACCAACGACATCACTACGAGGCTGTATGGGACTTCCGCTGCGCTGCACCT 407
QY 1276 CCGCCAGCTGCTGCGTGGCCCAACCAATGACTCCCATCAGCATTTATATCATCGAGCG 1335
Db 408 CCGCCAGCTGCTGCGTGGCCCAACCAATGACTCCCATCAGCATTTATATCATCGAGCG 467
QY 1336 GCAATAATGTGTCTACAGCAGTACGAGGACATGGTGGGAGTCGTGCGCTGCAGG 1395
Db 468 GCAATAATGTGTCTACAGCAGTATGAGGACATGGTGGGAGTCGTGCGCTGCAGG 527
QY 1396 TAG 1398
Db 528 TAG 530
```

```
RESULT 13
AAQ96215
ID AAQ96215 standard; DNA; 1345 BP.
XX
AC AAQ96215;
AT
DT 22-NOV-1995 (first entry)
XX
DE Human bone morphogenetic protein (BMP)-12 DNA.
XX
KW Bone morphogenetic protein; BMP-12; tendon; ligament; tendinitis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 138..1301
FT /*tag= a
FT mat_peptide 990..1301
FT /*tag= b
XX
PN W09516035-A.
XX
PD 15-JUN-1995.
XX
PF 06-DEC-1994; 94WO-US14030.
XX
PR 02-NOV-1994; 94US-0333576.
PR 07-DEC-1993; 93US-0164103.
PR 25-MAR-1994; 94US-0217780.
XX
PA (GENY ) GENETICS INST INC.
PA (HARD ) HARVARD COLLEGE.
XX
PI Celeste AJ, Melton DA, Rosen VA, Thomsen GH, Wolfman NM;
PI Wozney JM;
XX
XX
DR WPI; 1995-224320/29.
DR P-PSDB; AAR78734.
XX
XX
PT Bone morphogenetic proteins -12 and -13 and corresp. DNA - used in
PT compsn. for inducing tendon/ligament-like tissue formation
XX
PS Example; Page 73-75; 84pp; English.
XX
CC Oligo probe 3 (AAQ96214) designed on the basis of an amplified BMP-12
```

```
CC human DNA sequence (AAQ96212) was radioactively labelled and used to
CC screen a human genomic library. A positive clone was designated
CC lambda-HuG-48. This bacteriophage was deposited under ATCC 75625 on 7
CC Dec. 1993. The oligo hybridising region of this phage is localised
CC to a 3.2 kb BamHI fragment which was subcloned to PCR1-1#2 and
CC deposited under ATCC 69517 on 7 Dec. 1993. The partial DNA sequence
CC and derived AA sequence of the 3.2 kb DNA insert of the plasmid
CC subclone PCR1-1#2, derived from clone lambda-HuG-48 are shown in
CC AAQ96207 and AAR78729 respectively. Additional DNA sequence of the
CC 3.2kb BamHI insert of the plasmid PCR1-1#2 is given in AAQ96215.
CC All of AAQ96207 is contained in AAQ96215. As AAQ96215 is derived from
CC a genomic clone it is difficult to determine the boundary between
CC the 5' extent of coding sequence and the 3' limit of intervening
CC sequence.
XX
SQ Sequence 1345 BP; 205 A; 447 C; 475 G; 218 T; 0 other;
Query Match 18.2%; Score 259.2; DB 16; Length 1345;
Best Local Similarity 57.8%; Pred. No. 1.2e-37;
Matches 550; Conservative 0; Mismatches 378; Indels 24; Gaps 4;
QY 464 GACAGAAGTATTTTGTGATGTCTCATCTCTCAGACAAAGAGAGCTGGTGGGCGCG 523
Db 364 GCCAGAGCTTCCGTGTCGAGCTGTCCAGCCTTAACGACGACGAGAGTGGTGGTGCG 423
QY 524 AGCTGGCGCTCTTTTCGCCAGCGCCCTCAGCGCCCTGGGGGCCACACAGCGCGCGCTCC 583
Db 424 AGCTGGCGCTGTGCGCGGGATCTCCAGAGTC--GGGCCAGGACGCTGGACTTCTCC 481
QY 584 ACGTGACGCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTC 643
Db 482 -----GCCGTTGCTGCTGCTCCACGTGCGCGCGCGCGCGCGCGCGCTGC 534
QY 644 AGGGGCGCGCGCGCGCGCTGGGAAGTCTTCGACGTGTGGGAGGCGCTGGCCACCAGC 703
Db 535 TGTACTCGGGGCGAGCTGAGCCCTAGTCGGTTCAGCGCTGGGAGCGGCTTCGACGTGGCGG 594
QY 704 CTTGGAAGCAGCTGTGCTTGGAGCTGCGGGCGCATGGGCGGAGCTGGAGCGCGGGGAGG 763
Db 595 ACCCATGAGCGCCACCGCTGTAACCGCGCGCGCGCGCGCGCGCTTCTGCTTCTGCTGC 654
QY 764 CCGAGCGCGCGCGCGGGACCCAGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCTTGGGCT 823
Db 655 GCGCAGTGGCAGCGCGCGGTGCGCGAGCGCGTGGCACTGCGCGCGCTGGGCTGGCGTGC 714
QY 824 TCGGCGGAGGTCGCGCTTCCCGAGGAGCGCGCGCTTGGTGGTATTCACGATGCC 883
Db 715 CGGCGGAGGCGCTCTGCGGACAGAGCGCGCGCTGCTAGTCTCTCTCCCGCACGC 774
QY 884 AGCGCA---AGAACCTGTTTCGACAGATGCGGACGAGCTGGGCTCGCGCGAGGCTGC-- 938
Db 775 AGAGGAAGAGAGCTTATTCGGGAGATCGCGCCAGGCGCGCGCGCGCGCGCGCTC 834
QY 939 -----GGGCGCGCGCGCGCGCGCGAGGGGTGTTGGCGCGCGCGCGCGCGCGCGCG 988
Db 835 TGGCCTCAGAGCGCTGCGCGAGCAACCGGACCGCGCTGCGCAAGGCGAGTCATTG 894
QY 989 CGGATGCCAGGCTTGGCTGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1048
Db 895 GCGGCGCGCAGACGAGGAGGAGCGCGCTTGGCGGCGCGCGCGCGCGCGCGCGCGCG 954
QY 1049 GTCCGCTTGGCAAGCGCGCACGGAAGTCCAGGTACGCTGCAGCAAGAACCCCTGC 1108
Db 955 GGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 1014
QY 1109 ACCTGAACCTTCAAGGAGCTGGGCTGGGAGCTGGATTCGCGCGCGCGCGCGCGCGCGCG 1168
Db 1015 ACCTGAGCTTCAAGGAGCTGGGCTGGGAGCTGGATTCGCGCGCGCGCGCGCGCGCGCG 1074
QY 1169 CTTATCAGCTGCGAGGCTGATGCGACTTCCCGCTGCGCTGCGCACTGGAGCGCCACCAAC 1228
Db 1075 CGTACCACTGCGAGGCGCTTGGGAGCTTCCCTTTGGTGGCTGCGACCTCGAGCGCCACCAAC 1134
```



Db 1255 TCTAAGCAATACGAGCATGGTGGAGGCGCTGCGCTGCAGGTAGCG 1306

RESULT 15

AD18337  
ID AAD18337 standard; DNA; 1345 BP.

XX AAD18337;

XX 18-DEC-2001 (first entry)

DE Human bone morphogenic protein-12 (BMP-12) DNA #2.

XX Human; bone morphogenic protein; BMP-12; vulnery; antiinflammatory;  
KW analgesic; gene therapy; transforming growth factor-beta; TGF-beta;  
KW tissue formation; wound healing; tissue repair; ligament defect;  
KW carpal tunnel syndrome; tendonitis; ds.

XX Homo sapiens.

XX Key Location/Qualifiers  
FT CDS 138..1304

FT /tag= a  
FT /product= "Human BMP-12 protein"  
FT /note= "CDS does not include start codon"

FT /partial

FT sig\_peptide 138..989

FT /tag= b

FT mat\_peptide 990..1301

FT /tag= c

FT /product= "Mature BMP-12 protein"

XX US6284872-B1.

XX 04-SEP-2001.

XX 28-FEB-1997; 97US-0808324.

XX 22-DEC-1994; 94US-0362670.

XX 07-DEC-1993; 93US-0164103.

XX 25-MAR-1994; 94US-0217780.

XX 02-NOV-1994; 94US-0333576.

XX (GENY ) GENETICS INST INC.

PA (HARD ) HARVARD COLLEGE.

XX Celeste AJ, Wozney JM, Rosen VA, Wolfman NM, Thomsen GH;

PI Melton DA;

XX WPI; 2001-588978/66.

XX P-PSDB; AAE10986.

XX New chimeric DNAs, useful for treating tendonitis, carpal tunnel  
PT syndrome and other tendon and ligament defects, comprises DNA encoding  
PT propeptide linked to DNA encoding bone morphogenetic proteins (BMP)-12,  
PT BMP-13 or MP52 -

XX Example 1; Column 73-78; 42pp; English.

XX The invention relates to a chimeric DNA comprising a DNA sequence  
CC encoding a propeptide from a member of the transforming growth factor  
CC (TGF)-beta superfamily of proteins. This DNA is linked to a DNA sequence  
CC encoding an amino acid sequence encoding a mature polypeptide consisting  
CC of Bone Morphogenetic Protein (BMP)-12, BMP-13 or MP52 protein. The DNA  
CC sequences are useful for producing proteins which induce tendon/ligament  
CC like tissue formation, and for isolating and cloning further DNA  
CC sequences encoding BMP-12 related proteins with similar activity. The  
CC proteins are useful for the induction of tendon/ligament-like tissue  
CC formation, wound healing, ligament and other tissue repair, augmenting  
CC the activity of bone morphogenetic proteins, and for treating tendonitis,  
CC carpal tunnel syndrome and other tendon and ligament defects. The  
CC present sequence is a DNA encoding human BMP-12 protein.

XX

SQ Sequence 1345 BP; 205 A; 447 C; 475 G; 218 T; 0 other;

Query Match 18.2%; Score 259.2; DB 22; Length 1345;

Best Local Similarity 57.8%; Pred. No. 1.2e-37;

Matches 550; Conservative 0; Mismatches 378; Indels 24; Gaps 4;

QY 464 GACAGAAGTATTGTTGATGTGTCATGCTCTACACAAAGAGAGCTGGTGGCGCGG 523

DB 364 GCCAGAGCTTCCTGTCGACGTGTCACGCTTAACACGACAGAGGTGGTGGTGGCG 423

QY 524 AGTGGGCTCTTTTGCCAGGCGCCTCAGCGCCTTGGGGGCCACAGCGGCGCCCTCC 583

DB 424 AGTGGGCTCTTTTGCCAGGCGCCTCAGCGCCTTGGGGGCCACAGCGGCGCCCTCC 481

QY 584 AGTGGGCTCTTTTGCCAGGCGCCTCAGCGCCTTGGGGGCCACAGCGGCGCCCTCC 643

DB 482 -----GCCGTTGCTGCTGCTCCACGTGCGGCGCGCGCGCGCGCGCGCGCGCG 534

QY 644 AGGGGCG 703

DB 535 TGTACTCGCGGCGAGCTGAGCCCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGT 594

QY 704 CTGGAAGCAGCTGCTTGGAGCTGCGGGCCGCGATGGGCGGAGCTGGACGCGGGGAGG 763

DB 595 ACGCCATGAGCGGCCACCGCTGTAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 654

QY 764 CCGAGCG 823

DB 655 GCGCAGTGGCAGCG 714

QY 824 TCGCGCGGAGGTGCGGCGCTCCCGAGGAGCGGCGCGCGCGCGCGCGCGCGCGCGCG 883

DB 715 CGGGCGGAGGGGCTCTGCGGAGAGGAGCGCGCGGCTGCTAGTCTCTCTCCCGACGC 774

QY 884 AGCGCA---AGAACCTGTTGCGCAGAGATGCGGAGCAGCTGGGCTCGGCCGAGGCTGC-- 938

DB 775 AGAGGAAAGAGAGCTTATTCCGGGAGATCCGCGCCAGCGCGCGCGCGCGCGCGCGCTC 834

QY 939 -----GGGCG 988

DB 835 TGGCGCTCAGAGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 894

QY 989 CGGATGCCAGGCTTGGCTGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 1048

DB 895 GCGGCGCAGAGGAGGAGCGCGCTTGGCGGAGCGCGGAGCGGAGCGGAGCGGAGCGGCGG 954

QY 1049 GTGCGCATGGCAAGCGGCACGGCAAGAGTCCAGGCTAGCTGTCAGCAAGAGAGCCCTCG 1108

DB 955 GGGGCG 1014

QY 1109 AGGTGACTTCAAGAGCTGGGCTGGGAGCTGGGATGATTCGGCGCGCGCGCGCGCGCGCG 1168

DB 1015 AGTGGACTTCAAGAGCTGGGCTGGGAGCTGGGATGATTCGGCGCGCGCGCGCGCGCG 1074

QY 1169 CCTATCAGTGGAGGCTGTATGCGACTTCCCGCTGCGCTGCGCAGCTGGAGCGCCACCAAC 1228

DB 1075 CGTACACTGCGAGGCGCTTGGGACTTCCCTTGGCTTGGCACTTCGAGCGCCACCAAC 1134

QY 1229 AGCCCATCATCAGAGCTGATGAATCCATGAGACCGCGCGCGCGCGCGCGCGCGCGCG 1288

DB 1135 ATGCCATCATCAGAGCTGCTCAACTCCATGSCACAGACGCGCGCGCGCGCGCGCGCT 1194

QY 1289 GGTGCGCCACCAATGATCCCATCAGCATTCATATACATCGACGCGCGCGCGCGCGCG 1348

DB 1195 GTGTGCCAGCG 1254

QY 1349 TCTACAAGCAGTACGAGGAGCATGGTGGTGGAGTGGTGGTGGTGGTGGTGGTGGTGG 1400

DB 1255 TCTACAAGCAATACGAGGAGCATGGTGGTGGAGGCGCTGCGCGCTGCGAGGTAGCG 1306

Job time : 324 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2002, 02:52:20 ; Search time 51 Seconds  
(without alignments)  
857.670 Million cell updates/sec

Title: US-09-825-751A-20  
Perfect score: 2447  
Sequence: 1 MDTPRVLLSAVFLISFLWDL.....GNNVYKQYEDMVVESCGR 455  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1798	73.5	436	2 B55452	cartilage-derived
2	927.5	37.9	501	2 JC2347	growth/differentia
3	919.5	37.6	495	2 S43294	bone morphogenetic
4	918.5	37.5	501	2 A55452	cartilage-derived
5	691	28.2	125	2 S43295	bone morphogenetic
6	549	22.4	151	2 S43296	bone morphogenetic
7	470	19.2	398	2 JH0687	bone morphogenetic
8	464	19.0	398	2 JH0688	bone morphogenetic
9	453	18.5	405	2 I50608	bone morphogenetic
10	441.5	18.0	408	1 BMH04	bone morphogenetic
11	436.5	17.8	420	2 I49541	bone morphogenetic
12	434	17.7	393	2 S37073	bone morphogenetic
13	432.5	17.7	408	2 S38343	bone morphogenetic
14	431.5	17.6	396	1 BMH02	bone morphogenetic
15	429.5	17.6	461	2 S52408	SPDVR1 protein - s
16	426.5	17.4	408	2 JH0801	bone morphogenetic
17	423	17.3	394	2 S45355	bone morphogenetic
18	419.5	17.1	408	2 S58791	bone morphogenetic
19	418.5	17.1	452	2 I49542	bone morphogenetic
20	417.5	17.1	353	2 I50607	bone morphogenetic
21	416.5	17.0	402	2 A45056	osteogenic protein
22	416.5	17.0	454	1 BMH05	bone morphogenetic
23	413.5	16.9	401	2 JH0689	bone morphogenetic
24	408.5	16.7	400	2 A49147	bone morphogenetic
25	407	16.6	426	2 JH0690	bone morphogenetic
26	401.5	16.4	431	1 BMH07	bone morphogenetic
27	401	16.4	430	2 JQ1184	osteogenic protein
28	395.5	16.2	588	2 A26158	decapentaplegic pr
29	394.5	16.1	510	2 A54798	Vg-1-related prote

30	394.5	16.1	513	1 BMH06	bone morphogenetic
31	391	16.0	313	2 IS1284	bone morphogenetic
32	377	15.4	354	2 S29718	gene nodal protein
33	372	15.2	427	2 A40735	TGF beta homolog d
34	351	14.3	476	2 JC4646	bone morphogenetic
35	350.5	14.3	478	2 JC4838	bone morphogenetic
36	343	14.0	365	2 T43286	cet-1 protein - Ca
37	343	14.0	455	2 A43918	TGF-beta-related p
38	334	13.6	207	2 S37618	vgr protein - rat
39	325.5	13.3	472	1 BMH03	bone morphogenetic
40	324.5	13.3	360	2 A29619	Vgl embryonic grow
41	319.5	13.1	372	2 C39364	GDF-1 embryonic gr
42	318	13.0	426	1 B24248	inhibin beta-A cha
43	313.5	12.8	425	2 I47072	inhibin beta-A cha
44	310	12.7	424	1 WFGBA	inhibin beta-A cha
45	306.5	12.5	425	1 S50898	inhibin beta-A cha

ALIGNMENTS

RESULT 1

B55452 cartilage-derived morphogenetic protein 2 precursor - bovine (fragment)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 10-Feb-1995 #sequence\_revision 10-Feb-1995 #text\_change 26-May-2000  
C:Accession: B55452  
R:Chang, S.C.; Hoang, B.; Thomas, J.T.; Vukicevic, S.; Luyten, F.P.; Ryba, N.J.P.; Ko  
J. Biol. Chem. 269, 28227-28234, 1994  
A:Title: Cartilage-derived morphogenetic proteins. New members of the transforming gr  
A:Reference number: A55452; MUID:95050604; PMID:7961761  
A:Accession: B55452  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-436 <CHA>  
A:Cross-references: GB:U13661; NID:9632489; PIDN:AAA61416.1; PID:9632490  
C:Superfamily: inhibin

Query Match 73.5%; Score 1798; DB 2; Length 436;  
Best Local Similarity 80.8%; Pred. No. 8.1e-120;  
Matches 354; Conservative 18; Mismatches 48; Indels 18; Gaps 8;

QY	33	SSAELGSTRKMRSRKEGKQMRAPRDSADGRE---	GGEPPQPRQDEPR---	AQPPRAQEP	86
DB	2	ASAEISAKGMRTRKEGRMPRAPRENATAREPLDR	QEPPEPQRRPPQPEAREPP	61	
QY	87	GRGPVVPVPEYMLSIYRTYSIAEKLGINASFFOSS	SANTITTSFVDRGLDLSHTPLRRQ	146	
DB	62	GRGPLVPEYMLSIYRTYSIAEKLGINASFFOSSAN	TITSFVDRGLDLSHTPLRRQ	121	
QY	147	KYLFVYSMLSDKEELVGAELRLFRQAPSPAGPPAG	PLHVQLFPCLSPLLDARTLDPOG	206	
DB	122	KYLFVSTLSDEELVGAELRLFRQAPSPAGPPAGP	PLHVQLFPCLSPLLDARTLDPOG	179	
QY	207	APPAGVEFDVWGLRHQPMKQLCLLELRAAW-GEL	DAGEAEARARQPPPPDLRLSGF	265	
DB	180	APRGVEFDVWGLRHQPMKQLCLLELRAAWGEPGA	EAEARARQPPPPDLRLSGF	239	
QY	266	GRVRPQERALLVVTQRQKMLFAEMREQLGSA-EA	AGPGAGAGSGW-----PPPS	317	
DB	240	GRVTRPQERALLVVTQRQKMLFAEMREQLGSAE	AGPGAGAGSGW-----PPPS	299	
QY	318	GAPDAPFWLSPGRRRRRTAFASRHGKSRKSLRCS	KPLHVNFKELGWDWIAPLE	377	
DB	300	GTPDAGLWSPG-RRRTAFASRHGKSRKSLRCSKPL	HVNFKELGWDWIAPLE	358	
QY	378	YEAYHCEGVCDFPLRSHLEPTNHAIQTLMNSMDP	GSTPPSCCVPTKLTPIISLYIDAGN	437	
DB	359	YEAYHCEGVCDFPLRSHLEPTNHAIQTLMNSMDP	GSTPPSCCVPTKLTPIISLYIDAGN	418	
QY	438	NVYKQYEDMVVESCGR	455		
DB	419	NVYNEYEMVESCGR	436		

## RESULT 2

JC2347  
 growth/differentiation factor 5 - human  
 C:Species: Homo sapiens (man)  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 17-Mar-2000  
 C:Accession: JC2347  
 R:Hoetten, G.; Neidhardt, H.; Jacobowsky, B.; Pohl, J.  
 Biochem. Biophys. Res. Commun. 204, 646-652, 1994  
 A:Title: Cloning and expression of recombinant human growth/differentiation factor 5.  
 A:Reference number: JC2347; MUID:95071375; PMID:7980526

A:Accession: JC2347  
 A:Molecule type: DNA  
 A:Residues: 1-501 <HOE>  
 A:Cross-references: GB:X80915; NID:9671524; PIDN:CAA56874.1; PID:9671525  
 C:Genetics:  
 A:Gene: GDB:BMP9  
 A:Cross-references: GDB:433948  
 A:Transons: 211/1  
 A:Superfamily: inhibin  
 A:Keywords: glycoprotein  
 F:189/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:381-382/Cleavage site: Arg-Ala (unidentified proteinase) #status predicted

Query Match 37.9%; Score 927.5; DB 2; Length 501;  
 Best Local Similarity 44.7%; Pred. No. 3.6e-58;  
 Matches 209; Conservative 64; Mismatches 98; Indels 97; Gaps 16;

QY 38 GSTKMRSRKEGKQAPRDSADAGREGQPPQPRQDEPRAQO----- 79  
 DB 81 GOTGGTLQPKDEPKLP-----PRPG-GPEPKPGHPOTQRTARTVTPKGQPGGRAP 134  
 QY 80 -----PRAQEPGPGPRVPHYEMLSYRTSYTAELKGINASFFQ 119  
 DB 135 PRAGSVSPSFLKKAKEPGPPREPKEPPFPITPHEYMLSYRTLSADRGKGNSSVKL 194  
 QY 120 SSKSANTITSFVDRGLDLSHTPLRRQKYLFDVSMLSKDELRLFRQAPS----- 174  
 DB 195 ENGLANTITSFDKGGDDRGV-VKQKRVYFDISAL-EKDGGLLGLALRLRKPSDTAKP 252  
 QY 175 -APWGPAGPLHVQLFPCLS-----PLLLDARTLDPGAPAGWEVDFVWQGLRH-QPWKQ 228  
 DB 253 AAPGGGAAQL--KLSSCPGSRQAPALLDVRV--PGLDGSWEVDFDIWKLFRNKNSAQ 308  
 QY 229 LCLELRAANGELDAGEAEARAGPQPPDLRLSLGFGRRVPPQPERALLVFTSRQKN 288  
 DB 309 LCLELE-AW---ERGRA-----VDLRGLGFDRAARQVHEKALFLVFGTKKRD 352  
 DB 289 L-FAEMREOLGSAEAGPGAGAGSWPPSGAPDARPWLPSPGRRRTAFASRGKRRHG 347  
 DB 353 LFEFEIKARSGQDDKT-----VVEYLFQSO-RRKRRAPLATROGKRPS 393  
 QY 348 KKSRLRCKKPLHVNFKELGWDWIIAPLEYAEHCEGVCDPPLRSHLEPTNHAIQTILM 407  
 DB 394 KNLKARCSKALHVNFKDMGWDDWIIAPLEYAEHCEGLCEPPLRSHLEPTNHAIQTILM 453  
 QY 408 NSMDPGSTPPSCCVTKLTPISILYIDAGNNVYKQYEDMVMVESGCCR 455  
 DB 454 NSMDPESTPTCCVTRLSPISILFIDSANNVYKQYEDMVMVESGCCR 501

## RESULT 3

S43294  
 bone morphogenetic protein-related protein (GDF5) - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 20-Oct-1994 #sequence\_revision 10-Nov-1995 #text\_change 17-Mar-2000  
 C:Accession: S43294  
 R:Storm, E.F.; Huyh, T.V.; Copeland, N.G.; Jenkins, N.A.; Kingsley, D.M.; Lee, S.J.  
 Nature 368, 639-643, 1994  
 A:Title: Lmb alterations in brachypodism mice due to mutations in a new member of the  
 A:Reference number: S43294; MUID:94195427; PMID:8145850  
 A:Accession: S43294

A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-495 <STO>  
 A:Cross-references: GB:U08337; NID:9488461; PIDN:AAA18778.1; PID:9488462  
 C:Superfamily: inhibin

Query Match 37.6%; Score 919.5; DB 2; Length 495;  
 Best Local Similarity 43.5%; Pred. No. 1.3e-57;  
 Matches 209; Conservative 74; Mismatches 107; Indels 91; Gaps 18;

QY 21 PGFOQASISSSCSSAELGSGTGMRSRKEGKQAPRDSADAGREGQPPQPRQDE----- 74  
 DB 60 PGHHYGVGATNARAK-GSSGQTQAKDEPRKMPRSG-----GSETKPGPSSQTRQAAA 113  
 QY 75 ----PRAQEP-----RAQEP-----PGRGPRVPHYEMLSYRT 104  
 DB 114 RTVTGKQLPGKASSKAGSAPSSFLKKTREPTREPKEPPFPITPHEYMLSYRT 173  
 QY 105 YSIAEKLGINASFFOSKSNANTITSFVDRGLDLSHTPLRRQKYLFDVSMLSKDELVGA 164  
 DB 174 LSDADRGKGNSSVKLEAGLANTITSFDKGGDDRG-PAVRKQRYVFDISAL-EKDGLLGA 231  
 QY 165 ELRLFRQAP---SAPWGPAGPL-HVQLFPCLS-----PLLLDARTLDPGAPAGWEVDF 216  
 DB 232 ELRLRKKPLDVAKPAVPSSGSRVAQLKSSCPGSRQAPALLDVRV--PGLDGSWEVDF 289  
 QY 217 VWQGLRH-QPWKQKLCLELRAANGELDAGEAEARAGPQPPDLRLSLGFGRRVPPQER 275  
 DB 250 TWKLFNRKNSAQLCLELE-AW---ERGRA-----VDLRGLGFTARTARQVHEK 333  
 QY 276 ALLVYFTSRQKNL-FAEMREOLGSAEAGPGAGAGSWPPSGAPDARPWLPSPGRRR 334  
 DB 334 ALFLVFGTKKRDLEFEIKARSGQDDKT-----VVEYLFQSO-RRKR 374  
 QY 335 RTAFASRGKRRHGKSRKRLCKPLHVNFKELGWDWIIAPLEYAEHCEGVCDPPLRSH 394  
 DB 375 RAPLANRQGRKPSKLNKARCSKALHVNFKDMGWDDWIIAPLEYAEHCEGLCEPPLRSH 434  
 QY 395 LEPTNHAIQTILMSMDPGSTPPSCCVTKLTPISILYIDAGNNVYKQYEDMVMVESGCCR 454  
 DB 435 LEPTNHAIQTILMSMDPESTPTCCVTRLSPISILFIDSANNVYKQYEDMVMVESGCCR 494  
 QY 455 R 455  
 DB 495 R 495

## RESULT 4

A55452  
 cartilage-derived morphogenetic protein 1 precursor - human  
 C:Species: Homo sapiens (man)  
 C:Date: 10-Feb-1995 #sequence\_revision 10-Feb-1995 #text\_change 17-Mar-2000  
 C:Accession: A55452  
 R:Chang, S.C.; Hoang, B.; Thomas, J.T.; Vukicevic, S.; Luyten, F.P.; Ryba, N.J.P.; Ko  
 J. Biol. Chem. 269, 28227-28234, 1994  
 A:Title: Cartilage-derived morphogenetic proteins. New members of the transforming gr  
 A:Reference number: A55452; MUID:95050604; PMID:7961761  
 A:Accession: A55452  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-501 <CHA>  
 A:Cross-references: GB:U13660; NID:9600731; PID:9600732  
 C:Genetics:  
 A:Gene: GDB:CDMP1  
 A:Cross-references: GDB:438940  
 C:Superfamily: inhibin

Query Match 37.5%; Score 918.5; DB 2; Length 501;

Best Local Similarity 44.3%; Pred. No. 1.5e-57;  
 Matches 207; Conservative 64; Mismatches 101; Indels 95; Gaps 15;

QY 38 GSTKMRSRKEGKQAPRDSADAGREGQPPQPRQDEPRAQO----- 79  
 DB 81 GOTGGTLQPKDEPKLP-----PRPG-GPEPKPGHPOTQRTARTVTPKGQPGGRAP 134  
 QY 80 -----PRAQEPGPGPRVPHYEMLSYRTSYTAELKGINASFFQ 119  
 DB 135 PRAGSVSPSFLKKAKEPGPPREPKEPPFPITPHEYMLSYRTLSADRGKGNSSVKL 194  
 QY 120 SSKSANTITSFVDRGLDLSHTPLRRQKYLFDVSMLSKDELRLFRQAPS----- 174  
 DB 195 ENGLANTITSFDKGGDDRGV-VKQKRVYFDISAL-EKDGGLLGLALRLRKPSDTAKP 252  
 QY 175 -APWGPAGPLHVQLFPCLS-----PLLLDARTLDPGAPAGWEVDFVWQGLRH-QPWKQ 228  
 DB 253 AAPGGGAAQL--KLSSCPGSRQAPALLDVRV--PGLDGSWEVDFDIWKLFRNKNSAQ 308  
 QY 229 LCLELRAANGELDAGEAEARAGPQPPDLRLSLGFGRRVPPQPERALLVFTSRQKN 288  
 DB 309 LCLELE-AW---ERGRA-----VDLRGLGFDRAARQVHEKALFLVFGTKKRD 352  
 DB 289 L-FAEMREOLGSAEAGPGAGAGSWPPSGAPDARPWLPSPGRRRTAFASRGKRRHG 347  
 DB 353 LFEFEIKARSGQDDKT-----VVEYLFQSO-RRKRRAPLATROGKRPS 393  
 QY 348 KKSRLRCKKPLHVNFKELGWDWIIAPLEYAEHCEGVCDPPLRSHLEPTNHAIQTILM 407  
 DB 394 KNLKARCSKALHVNFKDMGWDDWIIAPLEYAEHCEGLCEPPLRSHLEPTNHAIQTILM 453  
 QY 408 NSMDPGSTPPSCCVTKLTPISILYIDAGNNVYKQYEDMVMVESGCCR 455  
 DB 454 NSMDPESTPTCCVTRLSPISILFIDSANNVYKQYEDMVMVESGCCR 501

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2002, 02:54:00 ; Search time 36 Seconds  
(without alignments)  
371.873 Million cell updates/sec

Title: US-09-825-751A-20  
Perfect score: 2447  
Sequence: 1 MDTPRVLLSAVFLISLWDL.....GNNVYKQYEDMVESGGR 455

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents, AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1747	71.4	321	1	US-08-362-670B-26
2	1747	71.4	321	3	US-08-333-576C-26
3	1747	71.4	321	4	US-08-808-324-26
4	1747	71.4	321	5	PCT-US94-14030A-26
5	1202.5	49.1	263	1	US-08-362-670B-32
6	1202.5	49.1	263	3	US-08-333-576C-32
7	1202.5	49.1	263	4	US-08-808-324-32
8	1202.5	49.1	263	5	PCT-US94-14030A-32
9	927.5	37.9	501	1	US-08-288-508C-2
10	919.5	37.6	495	1	US-08-455-559-10
11	919.5	37.6	495	5	US-09-145-060-10
12	919.5	37.6	495	5	PCT-US94-00657-10
13	917.5	37.5	401	3	US-08-289-222E-3
14	917.5	37.5	401	4	US-09-054-526B-3
15	765	31.3	388	1	US-08-362-670B-34
16	765	31.3	388	3	US-08-333-576C-34
17	765	31.3	388	4	US-08-808-324-34
18	765	31.3	388	5	PCT-US94-14030A-34
19	731	29.9	134	1	US-08-581-529B-6
20	731	29.9	134	3	US-09-097-616-6
21	731	29.9	134	5	PCT-US94-07762-6
22	691.5	28.3	294	1	US-08-362-670B-2
23	691.5	28.3	294	3	US-08-333-576C-2
24	691.5	28.3	294	4	US-08-808-324-2
25	691.5	28.3	294	5	PCT-US94-14030A-2
26	661	27.0	119	1	US-08-581-529B-7
27	661	27.0	119	3	US-09-097-616-7

28	661	27.0	119	5	PCT-US94-07762-7	Sequence 7, Appl
29	627	25.6	411	1	US-08-362-670B-28	Sequence 28, Appl
30	627	25.6	411	3	US-08-333-576C-28	Sequence 28, Appl
31	627	25.6	411	4	US-08-808-324-28	Sequence 28, Appl
32	627	25.6	411	5	PCT-US94-14030A-28	Sequence 28, Appl
33	618	25.3	240	1	US-08-362-670B-30	Sequence 30, Appl
34	618	25.3	240	3	US-08-333-576C-30	Sequence 30, Appl
35	618	25.3	240	4	US-08-808-324-30	Sequence 30, Appl
36	618	25.3	240	5	PCT-US94-14030A-30	Sequence 30, Appl
37	553	22.6	161	2	US-08-581-528A-6	Sequence 6, Appl
38	553	22.6	161	5	PCT-US94-07799-6	Sequence 6, Appl
39	549	22.4	119	1	US-08-455-559-13	Sequence 13, Appl
40	549	22.4	119	4	US-09-145-060-13	Sequence 13, Appl
41	549	22.4	119	5	PCT-US94-00657-13	Sequence 13, Appl
42	549	22.4	120	1	US-08-362-670B-4	Sequence 4, Appl
43	549	22.4	120	3	US-08-333-576C-4	Sequence 4, Appl
44	549	22.4	120	4	US-08-808-324-4	Sequence 4, Appl
45	549	22.4	120	5	PCT-US94-14030A-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1  
US-08-362-670B-26  
: Sequence 26, Application US/08362670B  
: Patent No. 5658882  
: GENERAL INFORMATION:  
: APPLICANT: Celeste, Anthony J.  
: APPLICANT: Wozney, John  
: APPLICANT: Rosen, Vicki A.  
: APPLICANT: Wolfman, Neil  
: APPLICANT: Thomsen, Gerald H.  
: APPLICANT: Melton, Douglas A.  
: TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS  
: NUMBER OF SEQUENCES: 37  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: GENETICS INSTITUTE, INC.  
: STREET: 87 Bridgepark Drive  
: CITY: Cambridge  
: STATE: Massachusetts  
: COUNTRY: USA  
: ZIP: 02140

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/362,670B  
FILING DATE: December 22, 1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Lazar, Steven R.  
REGISTRATION NUMBER: 32,618  
REFERENCE/DOCKET NUMBER: 5202-D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617 498-8260  
TELEFAX: 617 876-5851  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 321 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-362-670B-26

Query Match 71.4%; Score 1747; DB 1; Length 321;  
Best Local Similarity 100.0%; Pred. No. 4.3e-141;  
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 DLSHTPLRQRYLFDVSMLSKDELVGAEIRLFRQAPSAPWGPPAGPLHVQLFPCISPLL 196  
|||||

Db 3 DLSHTPLRRQKYLFDVSMLSKDELVLGAEALRLFRQAPSPAGPLHVQLFPCLSPLL 62  
QY 197 LDARTLDPOGAPPAGWEVFDVWQGLRHQHPWKOLCLLELRAANGELDAGEAARARQPOPPP 256  
Db 63 LDARTLDPOGAPPAGWEVFDVWQGLRHQHPWKOLCLLELRAANGELDAGEAARARQPOPPP 122  
QY 257 PDLRLSLGFGRRVRPPQERALLVVFTRSORKNLFAEMREQLGSAAAGPGAGAGSGWPPP 316  
Db 123 PDLRLSLGFGRRVRPPQERALLVVFTRSORKNLFAEMREQLGSAAAGPGAGAGSGWPPP 182  
QY 317 SGAPDARWLPSPGRRRRRTAFASRHKHKGKSLRCSKKPLHVNFKELGWDWIIAPL 376  
Db 183 SGAPDARWLPSPGRRRRRTAFASRHKHKGKSLRCSKKPLHVNFKELGWDWIIAPL 242  
QY 377 EYEAHCEGVCDFPLRSHLEPTNHAIQTLMNSMDPGSTPPSCCVPTKLTPIISILYIDAG 436  
Db 243 EYEAHCEGVCDFPLRSHLEPTNHAIQTLMNSMDPGSTPPSCCVPTKLTPIISILYIDAG 302  
QY 437 NNVYKQYEDMVVESCGR 455  
303 NNVYKQYEDMVVESCGR 321

## RESULT 2

US-08-333-576C-26  
; Sequence 26, Application US/08333576C  
; Patent No. 6027919  
; GENERAL INFORMATION:  
; APPLICANT: Celeste, Anthony J.  
; APPLICANT: Wozney, John  
; APPLICANT: Rosen, Vicki A.  
; APPLICANT: Wolfman, Neil  
; APPLICANT: Thomsen, Gerald H.  
; APPLICANT: Melton, Douglas A.  
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENETICS INSTITUTE, INC.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/333,576C  
; FILING DATE: No. 6027919ember 2, 1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lazar, Steven R.  
; REGISTRATION NUMBER: 32,618  
; REFERENCE/DOCKET NUMBER: 5202-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617 498-8260  
; TELEFAX: 617 876-5851  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 321 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-333-576C-26

Query Match 71.4%; Score 1747; DB 3; Length 321;  
Best Local Similarity 100.0%; Pred. No. 4.3e-141;  
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 DLSHTPLRRQKYLFDVSMLSKDELVLGAEALRLFRQAPSPAGPLHVQLFPCLSPLL 196  
|||||

Db 3 DLSHTPLRRQKYLFDVSMLSKDELVLGAEALRLFRQAPSPAGPLHVQLFPCLSPLL 62  
QY 197 LDARTLDPOGAPPAGWEVFDVWQGLRHQHPWKOLCLLELRAANGELDAGEAARARQPOPPP 256  
Db 63 LDARTLDPOGAPPAGWEVFDVWQGLRHQHPWKOLCLLELRAANGELDAGEAARARQPOPPP 122  
QY 257 PDLRLSLGFGRRVRPPQERALLVVFTRSORKNLFAEMREQLGSAAAGPGAGAGSGWPPP 316  
Db 123 PDLRLSLGFGRRVRPPQERALLVVFTRSORKNLFAEMREQLGSAAAGPGAGAGSGWPPP 182  
QY 317 SGAPDARWLPSPGRRRRRTAFASRHKHKGKSLRCSKKPLHVNFKELGWDWIIAPL 376  
Db 183 SGAPDARWLPSPGRRRRRTAFASRHKHKGKSLRCSKKPLHVNFKELGWDWIIAPL 242  
QY 377 EYEAHCEGVCDFPLRSHLEPTNHAIQTLMNSMDPGSTPPSCCVPTKLTPIISILYIDAG 436  
Db 243 EYEAHCEGVCDFPLRSHLEPTNHAIQTLMNSMDPGSTPPSCCVPTKLTPIISILYIDAG 302  
QY 437 NNVYKQYEDMVVESCGR 455  
303 NNVYKQYEDMVVESCGR 321

## RESULT 3

US-08-808-324-26  
; Sequence 26, Application US/08808324  
; Patent No. 6284872  
; GENERAL INFORMATION:  
; APPLICANT: Celeste, Anthony J.  
; APPLICANT: Wozney, John  
; APPLICANT: Rosen, Vicki A.  
; APPLICANT: Wolfman, Neil  
; APPLICANT: Thomsen, Gerald H.  
; APPLICANT: Melton, Douglas A.  
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENETICS INSTITUTE, INC.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/808,324  
; FILING DATE: Herewith  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lazar, Steven R.  
; REGISTRATION NUMBER: 32,618  
; REFERENCE/DOCKET NUMBER: 5202-D  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617 498-8260  
; TELEFAX: 617 876-5851  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 321 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-808-324-26

Query Match 71.4%; Score 1747; DB 4; Length 321;  
Best Local Similarity 100.0%; Pred. No. 4.3e-141;  
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 DLSHTPLRRQKYLFDVSMLSKDELVLGAEALRLFRQAPSPAGPLHVQLFPCLSPLL 196  
|||||